Macroeconomics and Information Frictions

Dissertation
submitted to the
Faculty of Business, Economics and Informatics
of the University of Zurich

to obtain the degree of
Doktor der Wirtschaftswissenschaften, Dr. oec.
(corresponds to Doctor of Philosophy, PhD)

presented by

GUANGYU PEI
from China

approved in July 2018 at the request of
Prof. Dr. Fabrizio Zilibotti
Prof. Dr. Nir Jaimovich

The Faculty of Business, Economics and Informatics of the University of Zurich hereby authorizes the printing of this dissertation, without indicating an opinion of the views expressed in the work.

Zurich, July 18, 2018

The Chairman of the Doctoral Board: Prof. Dr. Steven Ongena

Dedicated to My Son

Acknowledgements

First and foremost, I want to thank my advisor, with the most sincerity, Prof. Fabrizio Zilibotti for his continuous support in the past five years of my Ph.D. study in Zurich. It has been the great honor for me to have the opportunity to learn from him, not only in Economics, but also on how to contribute to the general society in life.

Secondly, I want to express my gratitude to Prof. George-Marios Angeletos, Prof. Massimo Marinacci and Prof. Nir Jaimovich for the inspiring comments and great supports for my researches. Also, I am grateful to my coauthors, Dr. Heng Chen and Dr. Yulei Luo, for the great experience they gave to me when we were working together. I also want to thank the faculty and my colleagues in Department of Economics, University of Zurich for all the great memory they gave me in the last five years.

Also, I would like to express my gratitude to the UBS Center for Economics in Society, University of Zurich for the generous scholarship that finances my Ph.D. study.

Last but not the least, I am really thankful to my wife Yang LI, my mother Xianjuan JIN and my father Yanji PEI. Their long-lasting support, patience and love are the foundations of my life and dedication to do academic researches.

Overview

Macroeconomic environment features complicated interactions among rational agents. Therefore, knowing what the economic fundamentals are and knowing what others know are essential for making economic decisions. Assuming information frictions in macroeconomic environments is not only the more reasonable assumption that depicts the real world but also turns out to have significant impacts on the economy both positively and normatively. The focus of this dissertation is the macroeconomics of information frictions (a) in shaping agents' information acquisition decisions, i.e., Chapter 1, (b) in creating business cycle fluctuations when agents are ambiguity averse, i.e., Chapter 2 and finally (c) in explaining the origins of financial sentiment when the economy features endogenous information production and transmission, i.e., Chapter 3.

Chapter 1 is the joint work with Heng Chen and Yulei Luo, *"Attention Misallocation, Social Welfare, and Policy Implications"*. We study how agents allocate attention between private and public signals to reduce the uncertainty about observation noises when coordination is an important concern. We find that attention allocation is non-monotone in the endowed attention capacity. In response to more attention capacity, agents may decrease their attention on or even ignore the more accurate signal. The desire to coordinate with each other may generate inefficient information acquisition behaviours. As a result, social welfare may decrease when they have more attention to process information when misallocation is severe. Finally, we derive sufficient and necessary conditions under which multiple equilibria emerge and study the implications of equilibrium multiplicity for macroeconomic policies. This paper has been published in *Journal of Economic Dynamics and Control, 2015, 59: 37-57.*

In Chapter 2: *"Ambiguity, Pessimism and Economic Fluctuations"*, I develop a novel theory of ambiguity-driven business cycles that contributes to explain the co-movements across market confidence, belief divergence and aggregate economy. I extend the RBC model with (a) aggregate demand externalities; (b) ambiguity averse agents with preferences represented by the smooth model of ambiguity axiomatized by Klibanoff et al. [2005,2009] and finally (c) incomplete information over ambiguous aggregate fundamentals. With the smooth model of ambiguity, ambiguity shock is formulated in a Bayesian fashion, namely shock to the variance of agents' prior belief over possible models. I highlight the dual impacts of this Bayesian formulation of ambiguity shock: a positive ambiguity shock makes all agents, who are ambiguity averse, behave as if they believe the aggregate fundamental is turning bad and becoming more volatile. When the economy features imperfect coordination due to incomplete information, dual impacts of a positive ambiguity shock translates into depressed belief over aggregate demand and the increased incentives to use private information both when making output decisions or output forecasts. The former maps into depressed market confidence and the latter maps into heightened belief divergence. And finally, aggregate output falls due to the increase in the economy-wide pessimism over aggregate demand. In combination, a positive ambiguity shock generates recession with depressed market confidence and heightened belief divergence. Quantitatively, ambiguity shock is shown to be capable of generating co-movements across real quantities together countercyclical belief divergence as measured by the cross-sectional dispersion in output forecast in SPF dataset. Also, the estimated time series of market confidence closely tracks Sentiment Index in Michigan Survey of Consumer. Therefore, I conclude that fluctuations in market confidence, belief divergence, and the aggregate economy are nothing more than the many shades of ambiguity shock.

Finally, in Chapter 3: *"Financial Sentiments and Coordinated Information Provision"*, I demonstrate that financial sentiments, which has been documented to create much of the volatilities in the financial market, may stem from the process of information production and transmission. Within a sender-receiver game that features information production by analysts and information transmission from analysts to investors, I demonstrate that, in addition to fundamental equilibrium where aggregate investment only responds to fundamentals, there exist sentiment equilibrium, in which, other than exogenous fundamentals, endogenous non-fundamental aggregate uncertainty, i.e., sentiment, affects aggregate investment. I derive the necessary and sufficient condition for the existence of sentiment equilibrium stating that as long as analysts care so much about reporting the level of aggregate investment K or have sufficiently strong incentive to align own report with the other analysts' reports, sentiment equilibrium arises. Finally, welfare implications of self-fulfilling sentiment fluctuations are discussed.

Contents

Acknowledgements i

Overview ii

1 Attention Misallocation, Social Welfare and Policy Implications 2

1.1 Introduction . . . 3

1.2 Related Literature . . . 6

1.3 The Model . . . 8

1.3.1 Players, Payoffs and Coordination . . . 8

1.3.2 Information Structure . . . 9

1.3.3 Equilibrium . . . 11

1.3.4 Characterization . . . 13

1.4 Attention Allocation . . . 15

1.4.1 Non-monotonic Attention Allocation . . . 16

1.4.2 The Role of Correlation . . . 19

1.5 "Too Much of a Good Thing": Social Welfare Analysis . . . 20

1.6 Policy Issues . . . 24

1.6.1 Generalization and Efficient Use of Information . . . 24

1.6.2 Transparency of Public Announcement . . . 27

1.7 Alternative Information Structures . . . 29

1.7.1 Observe the fundamental directly . . . 30

1.7.2 Observe two public signals . . . 30

1.8 Conclusion 31

Appendix of Chapter 1: Attention misallocation, social welfare and policy implication **33**

1.A Proofs 33

1.B Technical Appendix 41

2 Ambiguity, Pessimism and Economic Fluctuations **44**

2.1 Introduction 45

2.2 Related Literature 51

2.3 The Simple Model without Capital 54

2.3.1 Physical Environment, Shocks and Information Structure . 54

2.3.2 Preferences and Interim Belief Systems 58

2.3.3 Remarks and Interpretations 61

2.4 Equilibrium Characterization 62

2.4.1 Equilibrium Definition 62

2.4.2 Optimality Conditions 63

2.4.3 Joint Approximation of Allocation and Belief 65

2.5 Impacts of Ambiguity Shock 68

2.5.1 Game Theoretic Interpretation 69

2.5.2 Market Confidence, Belief Divergence and Aggregate Fluctuations 70

2.5.3 Discussion: incomplete information and ambiguity aversion 76

2.5.4 A summary 79

2.6 The Dynamic RBC Model: Quantitative Evaluation 79

2.6.1 Model Setup 79

2.6.2 Quantitative Methodology 85

2.6.3 Calibration 88

2.6.4 Business-Cycle Moments and Aggregate Co-movements . 89

2.6.5 Belief Divergence 92
2.6.6 Estimated Market Confidence v.s. Sentiment Index 93
2.7 Conclusion 94

Appendix of Chapter 2: Ambiguity, pessimism and economic fluctuations 96

2.A Derivations and Proofs 96
2.B Accuracy of the Approximation 107
2.C Labor Wedges 110
2.D Details for Quantitative Analysis 111
2.D.1 Equilibrium Conditions 111
2.D.2 Solution Method 112

3 Financial Sentiments and Coordinated Information Provision 119

3.1 Introduction 120
3.2 The Model 124
3.2.1 Agents, Actions and Payoffs 124
3.2.2 Information Structure 126
3.2.3 Equilibrium Definition 127
3.3 Equilibrium Characterization 128
3.3.1 The Complete Information Benchmark 128
3.3.2 Fundamental Equilibrium and Sentiment Equilibrium . . 130
3.4 Welfare Implications of Sentimental Fluctuations 136
3.5 Conclusion 138

Appendix of Chapter 3: Financial sentiments and coordinated information provision 140

Bibliography 147

List of Figures

1.1 The equilibrium uniqueness and multiplicity 14

1.2 The equilibrium share of attention 17

1.3 The non-monotonicity of social welfare. 22

1.4 Finite vs. infinite capacity. 22

1.5 Central bank transparency and social welfare. 28

1.A.1 Patterns of bounds ∇_0 and ∇_1. 35

2.1 Market confidence and belief divergence 45

2.1 Timeline for period t: Static 58

2.1 Impact of ambiguity shock: main mechanism 70

2.2 SPF forecasts of real GDP growth rate 74

2.1 Timeline for period t: dynamic 83

2.2 Impulse responses of ambiguity shock 91

2.3 Estimated times-Series vs empirical proxies: sentiment 94

2.B.1 Accuracy of the approximation 109

List of Tables

2.1 Model Parameters 89

2.2 Bandpass-Filtered Moments of Aggregate Variables 90

2.3 Bandpass-filtered Moments: Labor Wedges 92

2.4 Bandpass-Filtered Moments of Belief Divergence 93

Chapter 1

Attention Misallocation, Social Welfare and Policy Implications

This paper is co-authored with Heng Chen and Yulei Luo.[1] and is published as

Chen, H., Y. Luo, and G. Pei (2015). "Attention Misallocation, Social Welfare and Policy Implications." Journal of Economic Dynamics and Control 59: 37-57.

[1]We thank George-Marios Angeletos, Melody Lo, Jun Nie, Alessandro Pavan, Wing Suen, and Satoru Takahashi. We also thank seminar participants at Second Conference on Rational Inattention and Related Theories and 19th International Conference of Computing in Economics and Finance for their helpful comments. Luo thanks the Hong Kong General Research Fund (#HKU791913) for financial support.

1.1. Introduction

Coordination games with heterogeneity in information and complementarity in action have been widely applied to macroeconomic environments, financial markets and even collective actions. In this paper, we study a scenario, where agents that play such a coordination game have only limited capacity to process relevant information. Consequently, they have to allocate their capacity optimally among various information sources and then take actions based on the information they acquire. As Sims (2005) argue, this information processing constraint may have significant welfare implications for understanding the effects of policies that reveal public information and can be critical when evaluating the optimality of policies, e.g., the transparency of public announcements.[2]

To abstract from a specific market structure and retain tractability, we formalize our model in a "beauty contest" framework, as in Morris and Shin (2002), where the payoff for an individual depends on the distance of his action from an unobservable state and from the average action. To take the best action, agents must estimate the underlying state and forecast the average action of others. There are two correlated signals that reveal noisy information about the fundamental, and they can be observed if agents pay attention to them. One of the signals is private and contains idiosyncratic noise, and the other is public, can be potentially observed by all agents and contains common noise. The main point of departure of our model is to assume that agents cannot perfectly observe these signals because they possess a limited capacity to process information. Consequently, agents can only observe these signals with idiosyncratic observation noises.

This setting has captured the essence of several important macroeconomic environments and financial markets. For example, consider a price setting model with monopolistic competition in an informationally segregated island economy. Firms want to minimize the distance between their price and a target price that is the weighted average of an unknown state (e.g., log of nominal output) and the aggregate price. The representative firm on each island has access to public information about the nominal output (e.g., the central bank's guidance) and some island-specific information (e.g., local forecast). They can extract information about the state of the economy through those public and local sources, both of which reveal noisy information about the aggregate state,

[2] Sims (2005) argues that "rational inattention may have far-reaching implications for macroeconomics and monetary policy generally, once its implications are fully worked out. In the meantime, though, it may shed some light on transparency in monetary policy."

and are relevant for forecasting the aggregate price level. Constrained by a limited amount of capacity, firms need to decide on the allocation of their attention to process each of the noisy signals.

Another case in point is institutional investors in financial markets. They typically have access to information from a public source, e.g., announcements by the central bank or annual reports released by the firms whose securities are traded, and information from private sources, e.g., research reports on macro conditions or specific firms provided by analysts in research departments. Both types of sources reveal noisy information about the market fundamental. Investors need to allocate their limited attention optimally so that they can estimate the true fundamental and the aggregate action before making decisions.

Questions naturally arise. How do individuals allocate their attention between these two information sources? Under what conditions, do multiple equilibria emerge? Does social welfare necessarily increase when individuals possess more information-processing capacity?

In this model, a number of forces interact and shape agents' decisions: decreasing returns to attention, the relative accuracy of the public signal to the private signal, the coordination motive and the correlation between the two signals.[3] It is more attractive for agents to observe the signal with higher accuracy, because a signal of higher quality helps agents estimate the state more accurately. However, the force of diminishing returns to attention provides agents with incentives to diversify their attention and spend their capacity on both signals.[4] The coordination motive tilts the attention allocation decision toward the public signal, which better aligns their actions. Furthermore, a higher correlation across signals dampens the effect of diminishing returns to attention and amplifies the effect of the coordination motive.

The four factors combined may lead to the fact that agents allocate their attention in a non-monotone fashion, i.e., attention paid onto each signal may

[3] Sims (2010) argues that finite capacity can be elastic in response to a change in environment, given that the marginal cost of information processing is constant. In this case, inattentive agents are allowed to adjust optimal capacity in such a way that the marginal cost of information-processing for the problem at hand remains constant, which is consistent with the concept of "elastic" capacity proposed in Kahneman (1973). **?**) notes that the two assumptions, i.e., constant capacity and constant marginal cost of information processing, are observationally equivalent in the sense that they lead to the same model dynamics governed by the Kalman gain. In this study, for simplicity, we focus on the fixed capacity assumption and do not consider the effect of prior uncertainty on elastic capacity.

[4] "Diminishing returns to attention" refers to the fact that the marginal increase in the agent's welfare is decreasing as capacity increases. Luo (2008) and Luo and Young (2010) illustrate this property in partial equilibrium permanent income models with inattentive agents.

not necessarily increase in the total amount of capacity. For example, one intriguing scenario may arise, where agents can first focus on the relatively more precise private signal and then diversify their attention when capacity is higher; however, when there is a further increase in capacity, they may reduce their attention on or even ignore the private signal of higher quality and instead focus on the relatively imprecise public signal. We label this phenomenon "attention misallocation." Further, when the coordination motive or correlation is sufficiently high, the relative accuracy is not extreme and the amount of capacity is not very high, multiple equilibria can arise in this model.

We also find a number of distinct results on social welfare, i.e., the average distance between individual decisions and the underlying state. First, social welfare may decrease when individuals possess more capacity to process information. This result hinges on the fact that agents may "misallocate" their attention from a social perspective and the misallocation may become more severe in response to higher capacity. When there is an increase in capacity, agents can observe signals more clearly and better estimate the underlying state. However, given the desire to align their actions, they may decrease the attention paid to the private signal, even though it is relatively more precise, and coordinate even more attention on the less precise public signal. When agents take action, they assign a larger weight to the observation of the public signal, which exacerbates the "overreaction" to the public signal and causes a decrease in social welfare.[5]

Second, the limit case of this model is the world of Morris and Shin (2002), in which agents possess an infinite amount of capacity and can therefore perfectly observe both signals. However, strikingly, social welfare in the Morris-Shin world may be even lower than that in our model with capacity-limited agents. On the one hand, with a finite amount of capacity, agents have a less precise estimation of the fundamental than that in the Morris-Shin world. On the other hand, agents may endogenously pay little attention to the public signal and therefore reduce their reliance on it in their action which, to a certain degree, alleviates the overuse of the public signal. We show that the second effect can dominate.

Third, our model also sheds some light on the debate about the transparency of monetary policy. Morris and Shin (2002) show that social welfare can de-

[5]In the existing rational inattention literature, welfare typically increases in the capacity of processing information. For example, Luo (2008) shows that the welfare loss due to finite capacity decreases with channel capacity within a partial equilibrium permanent income model. Maćkowiak and Wiederholt (2011) obtains the same result in a general equilibrium business cycle model.

crease when the central bank delivers a clearer public announcement due to an overreaction to the public signal. Svensson (2006) questions the empirical relevancy of this result and argues that it only holds when public information is implausibly imprecise. We show that endogenous attention allocation can amplify the "overreaction," so that social welfare can decrease, even when the precision of the public signal is reasonably high.

Finally, our results also offer a new perspective on the literature covering the efficient use of information. Angeletos and Pavan (2007) show that equilibrium use and the efficient use of information coincide if and only if the social and private values of coordination are the same. However, once we allow for an endogenous information structure, i.e., attention allocation of inattentive agents, this relationship breaks down.

From a theoretical point of view, we also contribute to the literature of rational inattention by characterizing attention allocation between two *correlated* signals. This is technically challenging because, in this case, some capacity has to be "wasted" to learn the correlated part twice, and it is difficult to separate the amount of capacity that effectively reduces the observation noise of one signal from the other. We have bypassed this difficulty by using a transformation in which we define "effective capacity" so that such a separation is feasible. The standard information processing constraint for independent signals becomes a special case in our formulation.

The rest of the paper is organized as follows. Section 1.2 discusses the related literature. Section 1.3 characterizes a beauty contest model with rational inattentive agents and examines the equilibrium properties of the model. Section 1.4 explores the attention allocation decision via comparative statics. Section 1.5 studies the social welfare implications of limited attention. Section 1.6 addresses policy issues that are discussed in the literature. Section 1.7 contains a discussion on the alternative information structures. The last section concludes.

1.2. Related Literature

There are two approaches to modeling information acquisition in the related literature: "costly acquisition" with the cost of acquiring information being convex in the precision of signals and "rational inattention" with a fixed amount of attention being split among signals. These two approaches captures two distinct aspects of learning. With the former, it is increasingly costly for agents to find additional evidence for the truth; with the latter, agents are in a process of increasingly directed or refined search for an answer when they spend more capacity in learning.[6] In other words, the costly acquisition approach assumes that agents incur increasing marginal cost when they acquire more information, while the rational inattention approach assumes the marginal cost of information acquisition is decreasing.

Pioneering studies that adopt the costly acquisition approach examine the implications of information acquisition in coordination games, e.g., Hellwig and Veldkamp (2009), Myatt and Wallace (2012) and Yang (2015). Hellwig and Veldkamp (2009) show that there could exist multiple equilibria in their coordination model, because agents make a binary signal purchase decision and a public signal becomes more valuable when others also purchase it, due to strategic complementarity. Myatt and Wallace (2012) show that equilibrium is unique, even with complementarity in action, when agents can also choose the observation noises. In contrast, we follow Sims (2003) and assume that agents split a fixed amount of capacity on the signals to be observed. In our model, we demonstrate that multiplicity can arise, when the information acquisition is continuous and when learning takes the form of increasingly directed and refined search, captured by the rational inattention approach.

We intend to offer a welfare analysis of the coordination game played by rationally inattentive agents to study the effect that attention allocation has on social welfare. Our setup differs from those of recent contributions to the literature that explicitly deals with welfare-related issues. Maćkowiak and Wiederholt (2009, 2011) study how individuals or firms allocate their attention among two independent states when they set the price in a market-based economy or take collective actions. In our case, the two signals are correlated. We explicitly characterize the role of their correlation in optimal attention allocation and show that correlation is of critical importance for the multiplicity and uniqueness of equilibrium. For example, a high coordination motive does not necessarily give rise to multiple equilibria, unless the correlation between the two signals is sufficiently high.

This paper is closely related to the literature on the efficient use of information, e.g. Angeletos and Pavan (2006) and Colombo, Femminis, and Pavan (2012). The latter studies the interaction between the inefficient use and acquisition of information. In their model, agents pay to gain private information and can observe the public announcement precisely. In contrast, agents in our

[6] A capacity can be considered approximately the number of binary signals that partition states of the world, and moreover, the interpretation of each signals depends on its predecessors, see Veldkamp (2011) for elaboration.

model can observe neither of the signals perfectly. Unlike their setup, which has a unique equilibrium, the rational inattention assumption in our model gives rise to the possibilities of multiple equilibria and of one of the signals being endogenously ignored.

Our work also contributes to the growing literature on the social value of public information. Cornand and Heinemann (2008) consider an interesting setup in which only a fraction of the agents are allowed to observe the public signal. In our model, agents can endogenously choose to ignore either public or private information, or diversify their attention between both. Myatt and Wallace (2009) study this issue in a model with multiple information sources that differ in the degree of publicity. In our model, the publicity of public information is endogenous: the idiosyncratic observation noise is determined by the amount of attention paid.

This paper is also broadly related to the literature on information choice, attention allocation and asset allocation, which includes Peng (2005), Peng and Xiong (2006), Nieuwerburgh and Veldkamp (2009, 2010) and Mondria (2010). The framework adopted in these studies consists of multiple assets and a continuum of agents who face the information processing constraints.

1.3. The Model

1.3.1. Players, Payoffs and Coordination

The economy is occupied by a continuum of agents, indexed by $i \in (0,1)$. Each of them can choose an action, $a_i \in R$. In this economy, the fundamental state, θ, affects payoffs of agents. It is selected by nature but unknown to agents. Following Morris and Shin (2002), the payoff for agent i is specified by

$$u_i = -(a_i - \theta)^2 - \frac{\alpha}{1-\alpha}(L_i - \bar{L}), \tag{1.1}$$

where α is constant, such that $0 < \alpha < 1$, and $L_i = \int (a_j - a_i)^2 \mathrm{d}j$ and $\bar{L} = \int L_i \mathrm{d}i$.

When agent i takes action, two types of loss are incurred. The first component is measured by the distance between individual action and the uncertain state: agents would be better off if they were to choose an action closer to the fundamental. The second component is the distance between individual and average actions, which captures the idea that agents want to align their actions. A higher α implies that agents assign a larger weight to this strategic concern in their payoff structure and have a stronger incentive to coordinate.

1.3.2. Information Structure

Agents begin with some knowledge of the underlying state. Specifically, they share a common normal prior over θ,

$$\theta \sim N\left(\overline{\theta}, \sigma^2\right) \tag{1.2}$$

where $\overline{\theta}$ and σ^2 are the mean and variance of the prior distribution, respectively. Each agent can access two potentially observable signals, i.e., the private signal x_i and the public signal z, and the distribution is specified as follows,

$$s_i = \begin{pmatrix} x_i \\ z \end{pmatrix} \triangleq \begin{pmatrix} \theta + \varepsilon_{x_i} \\ \theta + \varepsilon_z \end{pmatrix}, \tag{1.3}$$

where $\varepsilon_{x_i} \sim N\left(0, \sigma_x^2\right)$ and $\varepsilon_z \sim N\left(0, \sigma_z^2\right)$ are independent of the true state θ. Note that ε_{x_i} is independently and identically distributed across agents while ε_z is common. The ex ante covariance matrix of s_i can be written as

$$\Sigma = \begin{pmatrix} \sigma^2 + \sigma_x^2 & \sigma^2 \\ \sigma^2 & \sigma^2 + \sigma_z^2 \end{pmatrix}. \tag{1.4}$$

The information structure described thus far resembles that in Morris and Shin (2002). The public signal can be interpreted as a public announcement made by the central bank or statistics released by the public agency. The private signal can be interpreted as information only accessible to individuals and not to the general public. Noise terms ε_{x_i} and ε_z can be interpreted as senders' noise contained in the signals, which cannot be reduced by paying attention to the signals. One implicit assumption is that agents cannot directly observe the fundamental and can obtain information only through analyzing the noisy signals about it.

Following Sims (2003), we assume that agents have a finite capacity to process available information, and that the reduction in uncertainty about the true signals is limited by finite entropy. Therefore, agents can only observe the noisy signals:

$$\hat{s}_i = \begin{pmatrix} \hat{x}_i \\ \hat{z}_i \end{pmatrix} = \begin{pmatrix} x_i \\ z \end{pmatrix} + \begin{pmatrix} \xi_{x_i} \\ \xi_{z_i} \end{pmatrix}, \tag{1.5}$$

where $(\xi_{x_i} \;\; \xi_{z_i})'$ are *observation noises*, which are independent of the true state and the sender noises, and are independently and identically distributed across agents. The presence of observation noises reflects the finite information processing capacity. Its co-variance matrix is given by

$$\Lambda = \begin{pmatrix} \omega_x^2 & 0 \\ 0 & \omega_z^2 \end{pmatrix} \tag{1.6}$$

where ω_x^2 and ω_z^2 are variances in the observation noises for private and public signals, respectively. Because the observation noises are idiosyncratic, noisy observation of the public signal, $\hat{z}_i$, is imperfectly correlated across the agents, whereas the observation of the private signal, $\hat{x}_i$, remains independent.

We define the posterior covariance matrix of s_i with $\Psi \equiv Var(s_i|\hat{s}_i)$, which can be determined using the following Gaussian updating formula,

$$\Psi = \Sigma - \Sigma\left(\Sigma + \Lambda\right)^{-1}\Sigma \text{ or } \Psi^{-1} = \Sigma^{-1} + \Lambda^{-1}. \tag{1.7}$$

We assume that each agent in this economy possesses a limited amount of capacity to process information. Specifically, each agent is assumed to face the following information-processing constraint:

$$\frac{1}{2}\ln\left(\frac{|\Sigma|}{|\Psi|}\right) \leq \kappa, \tag{1.8}$$

where $|\Sigma|$ and $|\Psi|$ denote the determinant of the prior covariance matrix of s_i and the corresponding posterior covariance matrix, respectively, and κ is positive and denote the individual's finite channel capacity. This constraint implies that the reduction in the uncertainty about the state gained from observing new signals is bounded from above by a finite capacity.

Given the specifications of (1.4), (1.6) and (1.7), the information processing constraint (1.8) can be rewritten as

$$\underbrace{\frac{1}{2}\ln\left(\frac{\omega_x^2 + \left(\sigma_x^2 + \sigma^2\right)\left(1-\rho^2\right)}{\omega_x^2}\right)}_{\hat{\kappa}_x} + \underbrace{\frac{1}{2}\ln\left(\frac{\omega_z^2 + \left(\sigma_z^2 + \sigma^2\right)\left(1-\rho^2\right)}{\omega_z^2}\right)}_{\hat{\kappa}_z} \leq \hat{\kappa}, \tag{1.9}$$

where ρ is the prior correlation across signals and $\hat{\kappa}$ is the effective capacity. They are defined by,

$$\rho \equiv Corr(x_i, z) = \sqrt{\frac{\sigma^2\sigma^2}{\left(\sigma_x^2 + \sigma^2\right)\left(\sigma_z^2 + \sigma^2\right)}}, \quad \hat{\kappa} \equiv \frac{1}{2}\ln\left(\exp\left(2\kappa\right) - \rho^2\left(\exp\left(2\kappa\right) - 1\right)\right).$$

Effective capacity is the amount of capacity used to reduce observation noises while a certain amount of capacity must be "wasted" to learn the correlated part of the two signals twice. Intuitively, for a fixed amount of κ, the higher is the correlation between signals and the lower is the effective capacity, $\hat{\kappa}$. If the signals are independent, i.e., $\rho = 0$, then $\kappa = \hat{\kappa}$.

The effective capacity spent on the private and public signals are denoted by $\hat{\kappa}_x$ and $\hat{\kappa}_z$, respectively. Naturally, we impose the following non-negativity restriction,

$$\hat{\kappa}_z \geq 0, \qquad \text{and} \qquad \hat{\kappa}_x \geq 0. \tag{1.10}$$

The variances in observation noises can be recovered from equation (1.9) as follows:

$$\omega_x^2 = \frac{\left(\sigma_x^2 + \sigma^2\right)\left(1 - \rho^2\right)}{\exp(2\hat{\kappa}_x) - 1}, \qquad \omega_z^2 = \frac{\left(\sigma_z^2 + \sigma^2\right)\left(1 - \rho^2\right)}{\exp(2\hat{\kappa}_z) - 1}. \tag{1.11}$$

If agents spend more effective capacity on observing a signal, then that signals' observation noise is smaller or it is clearer to the agents. In a limit case, agents possess an infinite amount of capacity and can therefore perfectly observe both signals, i.e., $\omega_x^2 = 0$ and $\omega_z^2 = 0$. To facilitate the characterization that follows, we define the relative accuracy of the public signal by

$$\nabla \equiv \sqrt{\frac{(\sigma_x^2 + \sigma^2)}{(\sigma_z^2 + \sigma^2)}},$$

and it can be readily verified that $\rho < \nabla < \frac{1}{\rho}$.

1.3.3. Equilibrium

This model environment can be considered as a two-stage game. In the first stage, nature draws the underlying state and agents make decisions on their attention allocation by optimally splitting the effective capacity between the signals to be observed. In the second stage, agents observe the realized signals and then take action.

We focus on a linear symmetric equilibrium in which all agents follow the same strategy in attention allocation and adopt a linear strategy in actions. Because the attention allocation is determined in the first stage, the heterogeneity in signal observations in the second stage does not affect their decision. Once agents decide their attention allocation, the variances in observation noises are also determined. The action strategy a_i in the second stage is linear in both the prior and observations on signals,

$$a_i = \Pi_{\theta,i}\bar{\theta} + \Pi_{x,i}\hat{x}_i + \Pi_{z,i}\hat{z}_i, \tag{1.12}$$

where $(\Pi_{\theta,i}, \Pi_{x,i}, \Pi_{z,i})$ are the weights assigned to the prior and observations.

We first solve the second-stage game, where the equilibrium remains unique. The solution to this game is the optimal weighting rule for any arbitrary allocation of attention in the first stage. Then, we solve for the optimal attention allocation in the first stage, taking the optimal weighting rule as given.

Given the linearity of the strategy and the normality of the information structure, we can show that an agent's action is a weighted average of the observations and their prior. That is,

$$\Pi_{\theta,i} + \Pi_{x,i} + \Pi_{z,i} = 1. \tag{1.13}$$

Individual i's payoff depends on the other agents' choices. Let the action strategy of the other agents be $(\hat{\kappa}_x, \hat{\kappa}_z, \Pi_\theta, \Pi_x, \Pi_z)$. The expected utility of individual i, $E[u_i]$, can be written as the sum of three components,

$$E[u_i] = -\underbrace{\frac{1}{1-\alpha}\left(\Pi_{\theta,i}^2\frac{1}{\phi_\theta} + \Pi_{x,i}^2\frac{1}{\phi_{x,i}} + \Pi_{z,i}^2\frac{1}{\phi_{z,i}}\right)}_{\mathcal{L}^\dagger} - \underbrace{\frac{\alpha}{1-\alpha}\left((\Pi_{z,i}-\Pi_z)^2\sigma_z^2 + (\Pi_{\theta,i}-\Pi_\theta)^2\sigma^2\right)}_{\mathcal{L}^\ddagger} + \mathcal{C}, \tag{1.14}$$

with

$$\phi_\theta = \frac{1}{(1-\alpha)\sigma^2}, \qquad \phi_{x,i} = \frac{1}{\sigma_x^2 + \omega_{x,i}^2}, \qquad \phi_{z,i} = \frac{1}{(1-\alpha)\sigma_z^2 + \omega_{z,i}^2}, \tag{1.15}$$

and

$$\mathcal{C} = \frac{-\alpha}{1-\alpha}\left[\left(\int a_j^2 dj - \bar{a}^2\right) - \bar{L}\right].$$

The first component of the utility function, $-\mathcal{L}^\dagger$, is the quadratic loss of agent i, when everybody uses the same weights on their signals in action strategy. It shows that both sender and observation noises are of importance for the expected loss. The second component, $-\mathcal{L}^\ddagger$, is another possible source of expected loss for agent i: utilizing a different action strategy from that used by the other agents. Obviously, when agent i adopts the same strategy, $\mathcal{L}^{\ddagger}$ becomes zero. Moreover, $\mathcal{C}$ is the effect of actions taken by other agents on agent i and it cannot be affected by her choice. Note that $\omega_{x,i}^2$ and $\omega_{z,i}^2$ are the variances of agent i's observation noises on the private and public signals, respectively, which are given by (1.11), and $\bar{a}$ is the aggregate action.

1.3.4. Characterization

Agent i maximizes $E[u_i]$ by choosing $\hat{\kappa}_{z,i}$ optimally.[7] To analyze the equilibrium allocation of attention, we study agent i's best response allocation strategy. We begin by defining the relative marginal return of attention on the public signal, which turns out to be convenient in the analysis of the main mechanisms. That is,

$$\gamma\left(\hat{\kappa}_{z,i}, \hat{\kappa}_z^*\right) \equiv \frac{\partial E\left[u_i\right]/\partial \hat{\kappa}_{z,i}}{\partial E\left[u_i\right]/\partial \hat{\kappa}_{x,i}}. \tag{1.16}$$

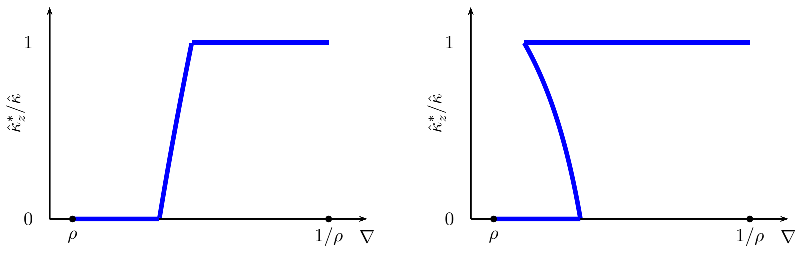
It is straightforward that $\gamma\left(\hat{\kappa}_{z,i}, \hat{\kappa}_z^*\right)$ measures the relative attractiveness of paying additional attention to the public signal, given others' attention allocation plan $\hat{\kappa}_z^*$. There are four forces that shape the attention allocation decision and thus affect γ. In the following sections, we fix the correlation between the two signals and discuss the role of the other three. We elaborate on the effect of the correlation in Section 1.4.2.

Lemma 1.1. $\gamma\left(\hat{\kappa}_{z,i}, \hat{\kappa}_z^*\right)$ decreases in $\hat{\kappa}_{z,i}$, increases in ∇, and increases in α.

First, the force of diminishing returns to attention takes hold. The more attention that is paid to the public signal, the less attractive it becomes. Second, agents prefer the signal with higher accuracy; that is, the higher its relative accuracy, the more attractive it is. Third, the coordination motive tilts agents' choice toward learning the public signal because they are rewarded in two ways when they spend more attention on the public signal: they are better informed about the underlying true state and their actions are better aligned. In other words, due to the coordination motive, the relative attractiveness of the public signal is magnified. The first part of Lemma 1.1 also implies that agent i increases her attention on the public signal if and only if $\gamma\left(\hat{\kappa}_{z,i}, \hat{\kappa}_z^*\right) > 1$, and decreases her attention if and only if $\gamma\left(\hat{\kappa}_{z,i}, \hat{\kappa}_z^*\right) < 1$.

In symmetric equilibrium, we impose the condition that $\hat{\kappa}_{z,i} = \hat{\kappa}_z^*$, and three

[7] Note that the optimal weighting rule $\left(\Pi_{z,i}^*, \Pi_{x,i}^*, \Pi_{\theta,i}^*\right)$ in the second stage is uniquely determined by the attention allocation plan $\left(\hat{\kappa}_{z,i}^*, \hat{\kappa}_{x,i}^*\right)$.



(a) $\alpha < 1 - \rho/\exp(\hat{\kappa})$. (b) $\alpha \geq 1 - \rho/\exp(\hat{\kappa})$.

Figure 1.1. *The equilibrium uniqueness and multiplicity.*

situations can arise. First, agents spend all of their attention on the public signal, where $\gamma(\hat{\kappa}_z^*, \hat{\kappa}_z^*) > 1$ and $\hat{\kappa}_z^* = \hat{\kappa}$. Second, they spend all of their attention on the private signal, where $\gamma(\hat{\kappa}_z^*, \hat{\kappa}_z^*) < 1$ and $\hat{\kappa}_z^* = 0$. Third, they split their attention between both signals, where $\gamma(\hat{\kappa}_z^*, \hat{\kappa}_z^*) = 1$ and $\hat{\kappa}_z^* \in [0, \hat{\kappa}]$. The following proposition offers the complete equilibrium characterization.

Proposition 1.1. *In linear symmetric equilibria, agents adopt the same attention allocation strategy, where the equilibrium attention allocation is such that*

$$\hat{\kappa}_z^* = \begin{cases} 0 & \text{if } \nabla \leq \nabla_0 \\ \tilde{\hat{\kappa}}_z & \text{if } \nabla \in (\min\{\nabla_0, \nabla_1\}, \max\{\nabla_0, \nabla_1\}) \\ \hat{\kappa} & \text{if } \nabla \geq \nabla_1 \end{cases} \tag{1.17}$$

$$\hat{\kappa}_x^* = \hat{\kappa} - \hat{\kappa}_z^*$$

where

$$\tilde{\hat{\kappa}}_z = \frac{1}{2}\hat{\kappa} + \ln\left(\frac{(1-\alpha)(1-\rho\nabla) + (\nabla-\rho)\exp(\hat{\kappa}) - (1-\rho^2)}{(1-\alpha)(1-\rho\nabla)\exp(\hat{\kappa}) + (\nabla-\rho) - \nabla(1-\rho^2)}\right) \tag{1.18}$$

and

$$\nabla_0 = \frac{\exp(\hat{\kappa})\rho + 1}{\exp(\hat{\kappa}) + \rho}, \qquad \nabla_1 = \frac{(1-\alpha)(\exp(2\hat{\kappa}) - 1) + (1-\rho^2)}{(1-\alpha)(\exp(2\hat{\kappa}) - 1)\rho + \exp(\hat{\kappa})(1-\rho^2)}. \tag{1.19}$$

There exist multiple equilibria, i.e., $\hat{\kappa}_z^* = \{0, \tilde{\hat{\kappa}}_z, \hat{\kappa}\}$, *if and only if*

$$\nabla_1 < \nabla < \nabla_0 \qquad \text{and} \qquad \alpha \geq 1 - \frac{\rho}{\exp(\hat{\kappa})}; \tag{1.20}$$

otherwise, the equilibrium attention allocation is unique.

When the relative accuracy is extreme, agents find it optimal to focus on only one of the signals; that is, for a fixed amount of capacity $\hat{\kappa}$ and a coordination motive α, if the relative accuracy is sufficiently high, i.e., $\max\{\nabla_1, \nabla_0\} \leq \nabla$, then agents choose to only observe the public signal. If the relative accuracy is sufficiently low, i.e., $\nabla \leq \min\{\nabla_1, \nabla_0\}$, then agents choose to observe the private signal only.

When the relative accuracy is not too extreme, this model can admit either multiple equilibria or a unique equilibrium.[8] See Figure 1.1. Multiple equilibria can arise when the coordination motive (or correlation) is sufficiently high and/or the total amount capacity is relatively low. First, the relative accuracy cannot be too extreme for the existence of multiple equilibria. When everybody focuses on the lower quality signal, agent i finds that the benefit of deviating and instead focusing on the relatively more precise signal is dominated by the cost of adopting a different strategy from other agents. Second, the coordination motive must be sufficiently large, such that when the strategic concern is strong, agents are more severely punished for deviating from the strategy adopted by other agents and therefore have less incentive to do so. Third, if the amount of capacity available is too large, then it is too costly for agent i to follow the others' strategy, conditional on the rest of the population coordinating on a "wrong" choice. In contrast, there is only a unique equilibrium if the coordination motive (or correlation) is not sufficiently strong or capacity is large.

In both cases, a symmetric equilibrium is formed if all agents choose the global minimizer of $\mathcal{L}^\dagger$, because both $\mathcal{L}^\dagger$ and $\mathcal{L}^\ddagger$ (defined in equation (1.14)) achieve global minimization and no individual has an incentive to deviate from it. We label it *strategic utility maximizing equilibrium*, because it generates the maximum of strategic utility, which is defined by $E[u_i^s] \equiv -\mathcal{L}^\dagger - \mathcal{L}^\ddagger$; that is the component, on which the choice of agents has an influence.

1.4. Attention Allocation

In this section, we analyze the comparative statics of the equilibrium attention allocation, not only because the issue of attention allocation itself is interesting but also because it provides building blocks for our examination of social wel-

[8]Technically, $\mathcal{L}^\dagger$ can be either quasi-concave or quasi-convex in $\hat{\kappa}_{z,i}$. The equilibrium is unique if and only if it is quasi-concave. In this model, the entropy is not a convex function of signal precision and that is why multiple equilibria can possibly emerge in this model. In contrast, with the costly acquisition approach, the cost function of noise reduction is usually assumed to be convex.

fare in Section 1.5, and policy issues in Section 1.6. We also highlight the role that the correlation between the public and private signals plays in attention allocation, as it is absent in most of the previous literature.

1.4.1. Non-monotonic Attention Allocation

For any relative accuracy and coordination motive, when the capacity is sufficiently large, the effect of diminishing returns eventually dominates, which leads to a diversified attention allocation. Because both signals can be extremely clear, the coordination motive and relative accuracy do not affect the attention allocation, with agents simply splitting a large amount of attention evenly between the two signals.

Proposition 1.2. *The share of effective capacity devoted to the public signal, $\hat{\kappa}_z^*/\hat{\kappa}$, converges to $1/2$, when capacity is sufficiently high.*

When the capacity is not large enough, the three forces characterized in Lemma 1.1 are intertwined and affect how the equilibrium attention allocation responds to an increase in capacity.

Proposition 1.3. *If the public signal is less accurate than the private signal, agents specialize in learning the latter, when the capacity is low. When the capacity increases, the share of effective capacity devoted to the public signal is monotonically increasing, if the private signal is very precise or the coordination motive is not so strong; otherwise, it is hump-shaped.*

In the proof of the proposition, we offer a complete characterization of this comparative statics. If the private signal is sufficiently accurate or the coordination motive is sufficiently low, it is never worthwhile to only observe the public signal, despite the effect of the coordination motive. The key trade-off here is between the effects of diminishing returns and relative accuracy, with the former eventually dominating the latter when the capacity is higher. Let $\hat{\kappa}_0$ be the threshold value of $\hat{\kappa}$, at which agents are indifferent about specialization in the private signal or diversification. In this case, when the capacity is higher than $\hat{\kappa}_0$, the share of attention devoted to the public signal monotonically increases in $\hat{\kappa}$. See Figure 1.2(a).

If the coordination motive is strong, its effect manifests in the non-monotonicity of the share of attention spent on the public signal. See Figure 1.2(b). When there is an increase in capacity, both the diminishing returns and the coordination motive have larger effects, and both forces tilt the choice of attention

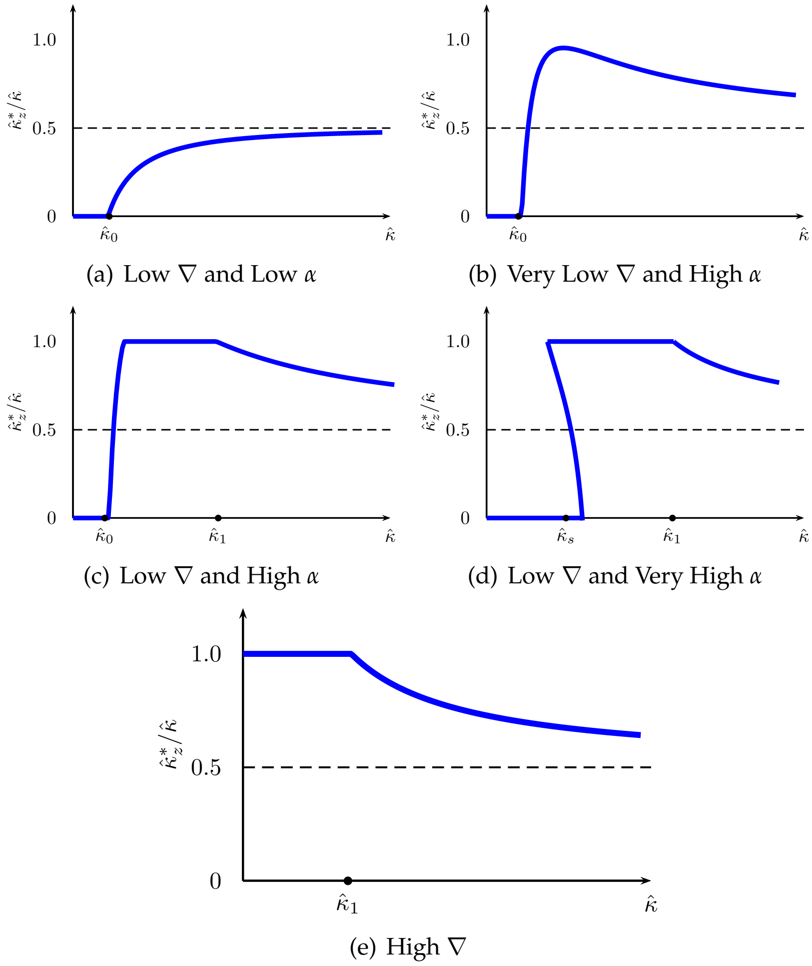


Figure 1.2. *The equilibrium share of attention allocated to the public signal is either monotonic or hump-shaped in effective capacity.*

allocation toward the public signal. Thus, there is a sharp increase in $\hat{\kappa}_z^*/\hat{\kappa}$. However, when agents allocate a predominant share of their attention to the public signal, the mechanism of diminishing returns to attention takes stronger effect and pushes agents to diversify. The effect of the coordination motive is eventually dominated and therefore, $\hat{\kappa}_z^*/\hat{\kappa}$ decreases in $\hat{\kappa}$.

If the precision of the private signal is close to that of the public one, the effect of the strong coordination motive can be so prominent that the share of attention on the public signal can reach 100%. See Figure 1.2(c). It is interesting to observe that in this case, agents' attention fans out, contracts inward, then fans out again. The number of signals that agents pick up does not monotonically increase in capacity.

Notably, $\hat{\kappa}_z^*/\hat{\kappa}$ being hump-shaped implies that, the absolute amount of attention paid to the more precise private signal can decrease (even to zero) when the total amount of capacity increases, as the result of a strong coordination motive. We formalize this mechanism with "attention misallocation," and demonstrate further how this mechanism critically affects the social welfare of this economy in Section 1.5.

Definition 1.1. *(Attention misallocation) Let j be the relatively more precise signal, i.e., $\sigma_j = \min\{\sigma_x, \sigma_z\}$ and $j \in \{x, z\}$. Attention misallocation arises, if the absolute amount of attention paid to the signal j in equilibrium decreases in response to an increase in total capacity, i.e., $\mathrm{d}\hat{\kappa}_j^*/\mathrm{d}\hat{\kappa} < 0$.*

Lemma 1.2. When the private signal is more precise than the public signal, the absolute amount of effective capacity allocated to observing the private signal can even decrease in the total amount of capacity, on the condition that $2\alpha + \rho > 1$. Specifically, $\mathrm{d}\hat{\kappa}_x^*/\mathrm{d}\hat{\kappa} < 0$.

When the coordination motive is very strong, all three equilibria can exist in the intermediate range of capacity. See Figure 1.2(d). Intuitively, this is the case in which none of the effects of relative accuracy, diminishing returns or coordination motive dominate the other two. Once other agents adopt one of the strategies, it is costly to deviate because the coordination motive is very high. Note that, in this case, the diversification equilibrium can never be the strategic utility maximizing equilibrium.[9] Therefore, if we focus on the strategic utility maximizing equilibrium, agents can shift their focus entirely from the private to the public signal when capacity crosses a cutoff value of $\hat{\kappa}_s$. The key trade-off here is between taking advantage of high accuracy and the desire for coordination.

In contrast, if the public signal is relatively more accurate, agents specialize in learning the public signal to take advantage of both higher accuracy and better coordination when the capacity is lower than $\hat{\kappa}_1$, i.e., the threshold value at which agents are indifferent between specialization or diversification. They eventually diversify, due to the effect of diminishing returns, and the equilibrium share of attention devoted to the public signal decreases monotonically. See Figure 1.2(e).

[9] When there exist multiple equilibria, $\mathcal{L}^\dagger$ is quasi-convex and the diversification allocation leads to a local minimum of $E[u_i^s]$.

1.4.2. The Role of Correlation

In this section, we turn to the role of correlation. The indirect effect of a change in correlation is straightforward. For any capacity κ, a higher correlation reduces the effective capacity available to agents, $\mathrm{d}\hat{\kappa}/\mathrm{d}\rho < 0$. Intuitively, because the two signals are correlated, observing both of them costs agents some capacity to learn the correlated part twice. The direct effect is characterized in the following proposition.

Proposition 1.4. *For any effective capacity, a higher correlation dampens the effect of diminishing returns to attention and amplifies the effect of the coordination motive. Specifically, (i) both $\hat{\kappa}_0$ and $\hat{\kappa}_1$ increase in ρ; (ii) $\hat{\kappa}_s$ decreases in ρ.*

First, for any amount of effective capacity available to agents, the observation noises are reduced more effectively when the correlation is higher. See equation (1.11). Because the two signals are correlated, knowing one of the signals helps reduce the other's observation noise. Therefore, agents have a stronger incentive to focus on one of the signals and the effect of diminishing returns is mitigated.

To demonstrate this, consider the case where only unique equilibrium exists for any capacity.[10] Due to the effect of diminishing returns, agents switch from specialization to diversification when the effective capacity is higher than $\hat{\kappa}_0$ for the case of $\nabla < 1$ and $\hat{\kappa}_1$ for the case of $\nabla > 1$. When the correlation is higher, both cutoff values increase. That is, agents find it worthwhile to diversify only when the effective capacity is at a higher level. In this respect, the effect of an increase in the correlation differs from that of a rise in the coordination motive α, which raises $\hat{\kappa}_1$ and lowers $\hat{\kappa}_0$.

Second, the correlation across signals also amplifies the effect of the coordination motive. Consider the case where the private signal is more precise. The rise in correlation entails a change in the trade-off between relative accuracy and coordination motive. In such a case, if agents spend more of their attention on the public signal, they estimate the underlying state less accurately but they can better align their actions. When the correlation between the private and public signals is higher, the two signals become more "substitutable," in terms of estimating the fundamental. Therefore, agents incur less welfare loss when they spend capacity on the less accurate public signal and they favor the public signal even more.

[10] As shown in Proposition 1.1, it is the case where $\alpha + \rho/\exp(\hat{\kappa}) < 1$.

There are three ways to see the effect of this mechanism. First, as shown in Proposition 1.1, for any effective capacity level $\hat{\kappa}$, and relative accuracy $\nabla \in [\nabla_1, \nabla_0]$, multiple equilibria emerge in this model when either α or ρ is sufficiently large. Second, when we consider the strategic utility maximizing equilibrium in this case, agents shift their focus from the private to the public signal at $\hat{\kappa} = \hat{\kappa}_s$. We observe that $\hat{\kappa}_s$ decreases in both ρ and α. Third, in Lemma 1.2, we show that, on the condition that either α or ρ is sufficiently high, attention misallocation can arise.

1.5. "Too Much of a Good Thing": Social Welfare Analysis

Social welfare is the average distance of individual actions in society from the fundamental. Agents benefit more from predicting the average opinion than other individuals, but it is a zero-sum game at the society level. In other words, the coordination motive only affects individual welfare and disappears at the society level. In this section, we analyze the comparative statics of social welfare by focusing on the strategic utility maximizing equilibrium.

The expected social welfare, $E\left[\mathcal{W}^s\left(\mathbf{a},\theta\right)\right]$, is a weighted average of $E\left[u_i^s\right]$, which is the objective expected utility maximized by agents, and the spillover effect, which is not considered by agents. The spillover receives a higher weight in social welfare if the coordination motive, α, is stronger.

$$\begin{aligned} E\left[\mathcal{W}^s\left(\mathbf{a},\theta\right)\right] &= -E\left[\int_i (a_i-\theta)^2\right] \\ &= \underbrace{(1-\alpha)\left[-\sigma^2\left(1+\frac{\phi_x}{\phi_\theta}+\frac{\phi_z}{\phi_\theta}\right)^{-1}\right]}_{E\left[u_i^s\right]} + \alpha \underbrace{\left[-\Pi_\theta^2\sigma^2 - \Pi_z^2\sigma_z^2\right]}_{\text{Spillover}}. \end{aligned}$$

The unintended spillover effect arises from agents' desire to align their actions and the fact that they do not consider the effects that their choices have on others. Agents make use of the common prior and their correlated noisy observations on the public signal, when they forecast the actions of others and choose their own actions. As both the prior and the public signal are noisy, the actions taken by agents may be anchored around commonly known but imprecise information. Therefore, the spillover contributes negatively to social welfare and its magnitude is determined by how precise the signals are, i.e., σ^2 and σ_z^2, and how much agents rely on them, i.e., Π_θ and Π_z.

Proposition 1.5. ***(Social Welfare and Capacity)*** *When the capacity to process in-*

formation increases, social welfare can decrease. Specifically, there may exist $\hat{\kappa}_a < \hat{\kappa}_b$, such that

$$E\left[\mathcal{W}^s(\hat{\kappa}_a)\right] > E\left[\mathcal{W}^s(\hat{\kappa}_b)\right].$$

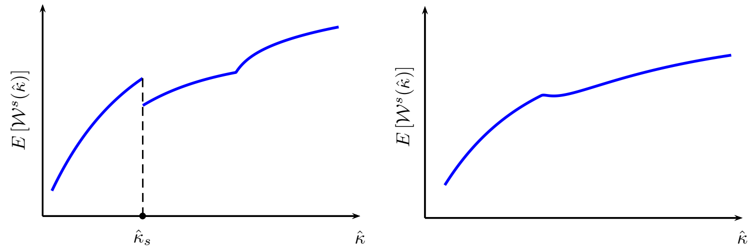
It is interesting to observe that a higher capacity to process information does not necessarily imply higher social welfare. On the one hand, $E\left[u_i^s|\hat{x}_i,\hat{z}_i\right]$, the part of welfare optimized by agents, always increases in capacity. On the other hand, the spillover can cause a decrease in social welfare when there is an increase in capacity. We know that spending more attention on the private signal reduces the spillover and enhances social welfare by lowering Π_z. However, recall the mechanism of attention misallocation shown in Lemma 1.2, agents may decrease the absolute amount of attention on the more accurate private signal when capacity increases, which results in a higher Π_z. Therefore, an increase in capacity can be detrimental to social welfare. In addition to this mechanism, holding constant $\hat{\kappa}_x^*$, a higher capacity implies that the amount of attention allocated to the public signal increases. Therefore, agents assign a larger weight to their observations on the public signal, $\hat{z}_i$, which is also socially costly.[11] In short, a strong coordination motive or a high correlation between signals can distort the allocation of attention so much that the spillover increases quickly in response to a higher capacity, which results in a decrease in overall social welfare.

To demonstrate this mechanism, we choose a set of parameters with high coordination motive (or high correlation) and intermediate relative accuracy, such that agents switch their attention entirely from the private signal to public signal at $\hat{\kappa}_s$, in the strategic utility maximizing equilibrium.[12] When $\hat{\kappa} = \hat{\kappa}_s$, agents are indifferent about only observing the private signal or the public signal; that is, $E\left[u_i^s|\hat{x}_i\right] = E\left[u_i^s|\hat{z}_i\right]$. In other words, $\phi_x = \phi_z$. For the same reason, Π_θ is the same in both cases. When $\hat{\kappa}$ increases from $\hat{\kappa}_s^-$ to $\hat{\kappa}_s^+$, $\hat{\kappa}_x^*$ decreases from $\hat{\kappa}_s$ to 0 and Π_z jumps from 0 to $1 - \Pi_\theta$, so that there is a discontinuous decrease in social welfare. Because $E\left[\mathcal{W}^s\right]$ monotonically increases in $\hat{\kappa}$, when $\hat{\kappa} < \hat{\kappa}_s$, there must exist $\hat{\kappa}_a$ and $\hat{\kappa}_b$ such that $E\left[\mathcal{W}^s(\hat{\kappa}_a)\right] > E\left[\mathcal{W}^s(\hat{\kappa}_b)\right]$ and $\hat{\kappa}_a < \hat{\kappa}_s < \hat{\kappa}_b$.

[11] We show that $\partial\Pi_\theta/\partial\hat{\kappa}_x^* = 0$; that is, the increase in ϕ_x must equal the decrease in ϕ_z when $\hat{\kappa}_x^*$ is optimally chosen. $\partial\Pi_\theta/\partial\hat{\kappa} < 0$, as holding $\hat{\kappa}_x^*$ constant, a higher capacity implies a higher ϕ_z and therefore, Π_θ must decrease. Intuitively, when the capacity is higher, agents rely more on their observation(s) and less on their prior knowledge.

[12] This situation arises, when $\alpha > 1 - \rho$ and $\tilde{\nabla} < \nabla < 1$, where $\tilde{\nabla}$ is defined in the proof of Proposition 1.3. Under this set of parameters, this model admits multiple equilibria and in this example, we focus on the change of social welfare in response to an change in capacity in a strategic utility maximizing equilibrium. However, this result does not rely on this particular case. In fact, the proof of Proposition 1.6 also implies Proposition 1.5. To establish Proposition 1.6, we focus on cases in which only a unique equilibrium exists.

See Figure 1.3(a).



(a) $\alpha = 0.7; \rho = 0.8; \nabla = 0.9135$ (b) $\alpha = 0.7; \rho = 0.8; \nabla = 0.8971$

Figure 1.3. *The non-monotonicity of social welfare.*

If the accuracy of the private signal is higher (or ∇ is lower), the absolute amount of attention paid to the private signal decreases gradually and the weight assigned to the observation on the public signal also increases gradually. Therefore, social welfare may decrease continuously in capacity. See Figure 1.3(b). The following equation summarizes the key mechanisms discussed above, where the sign of $+$ $(-)$ stands for a derivative being positive (negative).

$$\frac{\mathrm{d}E[\mathcal{W}^s]}{\mathrm{d}\hat{\kappa}} = (1-\alpha)\underbrace{\frac{\mathrm{d}E\left[u_i^s\right]}{\mathrm{d}\hat{\kappa}}}_{+} + \alpha\left[\underbrace{\frac{\mathrm{d}\left(-\Pi_\theta^2\sigma^2\right)}{\mathrm{d}\hat{\kappa}}}_{+} + \underbrace{\frac{\partial\left(-\Pi_z^2\sigma_z^2\right)}{\partial\hat{\kappa}_x^*}}_{+}\underbrace{\frac{\mathrm{d}\hat{\kappa}_x^*}{\mathrm{d}\hat{\kappa}}}_{+/-} + \underbrace{\frac{\partial\left(-\Pi_z^2\sigma_z^2\right)}{\partial\hat{\kappa}}}_{-}\right].$$

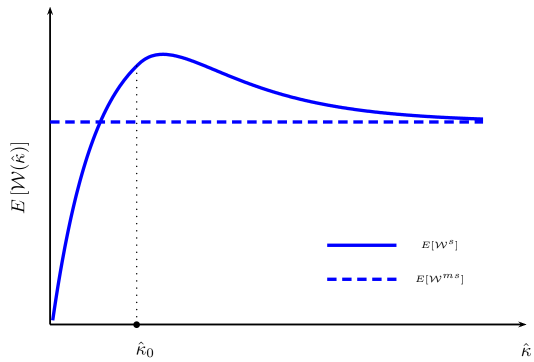


Figure 1.4. *Finite vs. infinite capacity.*

Proposition 1.6. ***(Too Much Capacity)*** *Social welfare can be higher when agents are endowed with a finite amount of capacity to process information than when they have an infinite amount of capacity. Specifically, there is a finite* $\hat{\kappa}'$*, such that*

$$E\left[\mathcal{W}^s(\hat{\kappa}')\right] > \lim_{\hat{\kappa}\to+\infty} E\left[\mathcal{W}^s(\hat{\kappa})\right] \equiv E\left[\mathcal{W}^{ms}\right].$$

This result is striking. When agents possess an infinite amount of capacity to process information, they can perfectly observe both signals. In this case, the model is identical to the Morris-Shin model, in which the social inefficiency is well understood, i.e., agents overreact to the public signal. Specifically, the weight agents assign to the public signal in their action is higher than that in their posterior belief, which is socially costly because the coordination motive driving the overreaction does not count in social welfare. Social welfare in the Morris-Shin model can be written as follows

$$E\left[\mathcal{W}^{ms}\right] = -\left[\frac{\phi_x^{ms} + \phi_z^{ms} + \frac{1}{(1-\alpha)}\phi_\theta}{\left(\phi_x^{ms} + \phi_z^{ms} + \phi_\theta\right)^2} + \frac{\Pi_z^{ms}\frac{\alpha}{(1-\alpha)}}{\phi_x^{ms} + \phi_z^{ms} + \phi_\theta}\right]\frac{1}{\sigma^4},$$

where variables with superscript ms are counterparts in the Morris-Shin model.

In our case, capacity-limited agents cannot clearly observe signals; thus, their estimation of the underlying state is less accurate than that when they have an infinite amount of capacity. However, agents may endogenously choose to spend very little attention on observing the public signal, as a result, the total amount of noise in the observation, $\hat{z}_i$, becomes very large. Therefore, they rely on it much less when they take actions; that is, the weight that it is assigned, Π_z, can be lower than Π_z^{ms}. A lower level of capacity can actually be welfare enhancing, because it does, to some extent, correct the inefficient use of public information. If the second effect dominates the first, social welfare can be higher than that in the Morris-Shin model. See Figure 1.4.

We demonstrate the two opposing effects with the following simple case. Let the total amount of capacity in our model be $\hat{\kappa}_0$. We choose a set of parameter such that agents are indifferent between specialization in the private signal or diversification; that is, they endogenously ignore the public signal, or, $\phi_z = 0$.[13] We write social welfare as follows

$$E\left[\mathcal{W}^{s}\right] = -\left[\frac{\phi_x + \frac{1}{(1-\alpha)}\phi_\theta}{\left(\phi_x + \phi_\theta\right)^2}\right]\frac{1}{\sigma^4}.$$

Given the finite capacity, agents cannot perfectly observe the private signal and thus the precision of observation of the private signal is smaller; that is, $\phi_x < \phi_x^{ms}$. Moreover, in the Morris-Shin model, the public signal is also informative and enhances the estimation of the underlying state, which results in

[13] Section 1.4 demonstrates that such a $\hat{\kappa}_0$ exists, unless both $\alpha + \rho > 1$ and $\nabla > \tilde{\nabla}$ hold, where $\tilde{\nabla}$ is defined in the proof of Proposition 1.3.

$\phi_x < \phi_x^{ms} + \phi_z^{ms}$. Intuitively, agents with finite capacity are always worse off in terms of estimating the underlying state. It always holds that

$$-\left[\frac{\phi_x + \frac{1}{(1-\alpha)}\phi_\theta}{(\phi_x + \phi_\theta)^2}\right]\frac{1}{\sigma^4} < -\left[\frac{\phi_x^{ms} + \phi_z^{ms} + \frac{1}{(1-\alpha)}\phi_\theta}{(\phi_x^{ms} + \phi_z^{ms} + \phi_\theta)^2}\right]\frac{1}{\sigma^4}.$$

The second term in $E[\mathcal{W}^{ms}]$ shows the additional welfare loss caused by overusing the public signal in the Morris-Shin model. There is no overuse of the public signal in this particular finite-capacity case in that $\Pi_z = 0$, because $\phi_z = 0$. The socially costly overreaction to the public signal does not exist in this case

$$0 > -\left[\frac{\Pi_z^{ms}\frac{\alpha}{(1-\alpha)}}{\phi_x^{ms} + \phi_z^{ms} + \phi_\theta}\right]\frac{1}{\sigma^4}.$$

When social inefficiency is high in the Morris-Shin model, the welfare loss due to "overreaction" to the public signal can be so large that the gain from a better estimation of the fundamental is dominated.

Given that the capacity can be "too much," is it possible for agents to voluntarily burn some capacity to achieve higher welfare? The answer is no. That is because, if everyone else collectively discards some of their capacity, individual i can increase her welfare by fully using all of her capacity to enhance the estimation of the fundamental in the first stage and adopting the same action strategy in the second stage to avoid being "punished" for using a different strategy.[14]

1.6. Policy Issues

In previous sections, we have fully characterized the optimal attention allocation and explored its implications for social welfare via comparative statics. In this section, we discuss two welfare-related issues to shed light on how policy prescriptions in the literature can be amended, considering that agents are capacity-constrained.

1.6.1. Generalization and Efficient Use of Information

Angeletos and Pavan (2007) offer a flexible efficiency benchmark to assess the welfare properties of a general class of games where the social value of coordination may be higher or lower than the private one. In this section, we demon-

[14] This argument can be formalized and its proof is available on request.

strate that our key results on social welfare can continue to hold in that generalized environment. Moreover, we also explain how the equilibrium multiplicity generated by our model mechanism would affect policy prescriptions offered in their work.

To accommodate this analysis, we enrich the payoff structure in this model by following Angeletos and Pavan (2007), where the utility function is specified by a general linear quadratic function, $u = U(k, K, \sigma_k, \theta)$, where k and K are individual and aggregate actions, respectively; $\sigma_k \equiv \int_i (k-K)^2 \, \mathrm{d}i$ is the action dispersion. In fact, we can write the utility function as follows,

$$\begin{aligned} u \equiv & \underbrace{-(k_i-\theta)^2 - r(k_i-K)^2 + r\sigma_k^2}_{I} \\ & + \underbrace{\left(\frac{U_{\sigma\sigma}}{2} - r\right)\sigma_k^2}_{II} + \underbrace{\left(\frac{U_{KK}}{2} + r\right)K^2}_{III} + \underbrace{(U_{kK} - 2r)\,kK}_{IV} \\ & + \underbrace{\left(\frac{U_{\theta\theta}}{2} + 1\right)\theta^2}_{V} + \underbrace{\left(\frac{U_{kk}}{2} + 1 + r\right)k^2}_{VI} + (U_{k\theta} - 2)\,k\theta + U_{K\theta}K\theta. \end{aligned}$$

where r is a positive constant. Our model is a special case, where we have $U_{\sigma\sigma} = 2r$, $U_{KK} = -2r$, $U_{kK} = 2r$, $U_{\theta\theta} = -2$, $U_{kk} = -2(1+r)$, $U_{k\theta} = 2$ and $U_{K\theta} = 0$.[15] Part I is the standard beauty contest utility specification, by letting $\alpha = \frac{r}{1-r}$.

The private value of coordination, or how much agents care about aligning their actions, is measured by α. The socially optimal degree of coordination is represented by α^* and it is the weight that the social planner would assign to the aggregate action in its best response. In the general case, they can be characterized as follows,

$$\alpha^* = 1 - \frac{U_{kk} + 2U_{kK} + U_{KK}}{U_{kk} + U_{\sigma\sigma}} \qquad \alpha = 1 - \frac{U_{kk} + U_{kK}}{U_{kk}}. \tag{1.21}$$

Note that the beauty contest game is a special case with $\alpha^* = 0$ (i.e., the social planner does not value coordination) and with $\alpha > 0$ (i.e., individuals care about aligning their actions). Efficient attention allocation and efficient use of information are derived by solving a social planner problem while respecting the information processing constraint (1.9).

First, in the benchmark case, we have demonstrated in Proposition 1.5 and

[15] We restrict our attention to the case where there is no inefficiency under complete information. That is, we impose a restriction on $U_{k\theta}$ and $U_{K\theta}$, which is standard as in Angeletos and Pavan (2007).

Lemma 1.2 that social welfare may decrease in response to an increase in capacity due to attention misallocation, on condition that coordination motive and/or correlation between signals are sufficiently strong. Alternatively, this result can be interpreted as follows. The social planner does not value coordination among agents at all in the beauty contest case. When the private value of coordination perceived by the agents is high, the equilibrium allocation may deviate a lot from the social optimum. Therefore, a further increase in capacity can result in an even severer attention misallocation, which may cause a reduction in social welfare.

In this general case, this mechanism can still produce the same effect, that social welfare decreases in capacity, when there is a large enough discrepancy between the value of coordination perceived by individuals, α, and the central planner, α^*. To illustrate this point, consider the following scenario. Suppose the payoff function is parametrized such that $\alpha = \alpha^*$. A positive $U_{\sigma\sigma}$ implies that the action dispersion has positive externality on individuals' payoffs. An increase in $U_{\sigma\sigma}$ reduces the social value of coordination, α^*.[16] Therefore, the central planner prefers a higher action dispersion and allocating even less attention to the public signal and more attention to the private signal. However, the equilibrium value of coordination is not affected by the change in $U_{\sigma\sigma}$. When $U_{\sigma\sigma}$ is large enough, the attention misallocation, due to the discrepancy in equilibrium and efficient degree of coordination, can intensify and cause a decrease in social welfare when capacity increases.[17]

Second, one of the key insights in Angeletos and Pavan (2007) with the exogenous information structure is that the equilibrium use of information is efficient if and only if the social and private values of coordination coincide. However, this result may not hold once we allow for an endogenous information structure. Further, in this model, even though attention allocation is efficient, it does not necessarily lead to an efficient use of information.

In contrast to Angeletos and Pavan (2007), we argue that even when the central planner corrects the coordination incentives of agents to the socially optimal level with a tax policy, the equilibrium attention allocation and the use of information may still be not socially optimal. The key to understanding this argument is to recall that multiple equilibria may arise. Consider the case where the socially optimal degree of coordination α^* is higher than the private value of coordination α and it is so high that there exist multiple solutions in the cen-

[16]Note that both $U_{kk} + 2U_{kK} + U_{KK}$ and $U_{kk} + U_{\sigma\sigma}$ are negative in this case.

[17]Similar arguments can be applied to other cases and detailed analysis is available on request.

tral planner's social welfare optimization problem. It is obvious that the central planner picks the solution that gives rise to the highest social welfare. In a decentralized economy, with a policy similar to that proposed in Angeletos and Pavan (2007), the central planner can incentivize agents to value the coordination as much as it does, but it is still undetermined on which equilibrium agents coordinate. The planner needs another set of tools that help direct agents to coordinate on the social welfare maximizing equilibrium.

Further, in this framework, the equilibrium use of information can still be inefficient, even though the equilibrium attention allocation is efficient. To see this, we assume that $\alpha > \alpha^* > 0$. The social planner also values the coordination and may dictate that agents focus on the public signal and ignore the private one. Under the same conditions, individuals could also choose exactly the same attention allocation. In other words, the attention allocation is socially optimal. However, as the private and social values of coordination differ, in the second stage, individuals would assign a higher weight in their action strategy to the common prior than would the social planner. To understand this, recall the fact that the common prior serves a "free public signal," which does not require any attention, and observations on the public signal are imperfectly correlated across agents in this economy due to the idiosyncratic observation noises. Therefore, the equilibrium use of information is still not socially optimal.

1.6.2. Transparency of Public Announcement

This study adds another dimension to the debate about central bank transparency. Unlike the common presumption that higher transparency is always beneficial, Morris and Shin (2002) show that it may be detrimental to social welfare when the central bank delivers clearer public announcements. In their beauty contest model with an exogenous information structure, an increase in the precision of public information entails two opposing effects. On the one hand, it allows agents to better estimate the underlying fundamental. On the other hand, it also increases agents' reliance on the noisy public information in their actions, which is socially costly. Both effects are enlarged as the precision of public information increases. Morris and Shin (2002) show that social welfare is U-shaped, such that when the precision of the public signal is exceedingly low, the second effect dominates. Specifically, social welfare decreases in

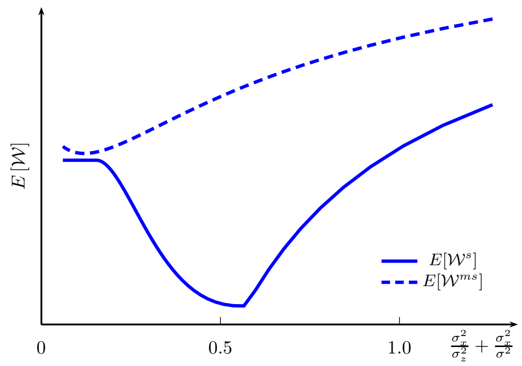


Figure 1.5. *Central bank transparency and social welfare.*

its precision if and only if

$$\frac{\sigma_x^2}{\sigma_z^2} + \frac{\sigma_x^2}{\sigma^2} < (2\alpha - 1)(1 - \alpha). \tag{1.22}$$

Therefore, it may be socially desirable to withhold public information.

One important critique of this argument is Svensson (2006), that questions its empirical relevancy and stresses that it can hold only when public information is implausibly imprecise.[18] See the dashed line in Figure 1.5 where social welfare is plotted against the left side of equation (1.22), holding σ_x^2 and σ^2 fixed.[19]

However, we argue that the precision of the public signal needs not necessarily to be exceedingly low to generate a decline in social welfare, when we allow for endogenous attention allocation. When agents can decide to which information source they pay their attention, the precision of each signal that they observe becomes endogenous in that it not only depends on variances in sender noises, but also on those of observation noises, which are chosen by agents.

To illustrate this, we plot social welfare in our model with the solid line in Figure 1.5.[20] When the precision of the public signal is very low, agents ignore it and focus on the private signal. Therefore, a marginal increase in the precision of the public signal does not affect social welfare. When the precision of the public signal is sufficiently high, agents diversify their attention. An increase

[18] Even the maximum of the right side of (1.22) is a very small number, which implies that σ_z^2 must be sufficiently large for this condition to hold.

[19] In this numerical example, $\alpha = 0.7$ and $\sigma_x = 0.1$. σ is normalized to unit.

[20] In this numerical example, the capacity available to agents is $\kappa = 3.2$ bits and all other parameters are the same as those for computing the counterpart in the Morris-Shin model. With this set of parameters, there is a unique equilibrium.

in its precision leads to a higher reliance on the public signal in their action, as in Morris and Shin (2002). In addition, agents also direct a larger proportion of their attention toward the public signal in response to a higher precision. This additional mechanism reinforces the previous one and both contribute to the decline in social welfare. As a result, social welfare still decreases, even when the precision of the public signal is reasonably large.[21]

1.7. Alternative Information Structures

Following Hellwig and Veldkamp (2009) and Myatt and Wallace (2012), we have made two implicit assumptions about the way in which agents learn in the benchmark model. First, agents cannot directly observe the fundamental, but can learn about it through analyzing noisy signals. In other words, agents can only obtain noisy observations about the underlying signals that are informative about the fundamental.[22] That assumption is realistic in settings in which information acquisition takes the form of the assimilation of the available information and agents face a constraint to information transmission and comprehension. Second, the source of noisy information can be both public and private.[23] That assumption captures the scenario that agents may have access to information sources of a different nature, i.e., economy-wide information that is commonly accessible and local information that is conditionally uncorrelated.[24] This set of assumptions, although common in the literature, is important for our results. The following discussion outlines how our results would change if we have allowed for alternative information structures.

[21]When the precision of the public signal is high enough, agents pay all their attention to the public signal. In this situation, an increase in the precision of the public signal is always welfare-enhancing.

[22]The rational inattention model characterized by Myatt and Wallace (2012) allows agents to learn one noisy signal about the fundamental that contains exogenous sender's noise. They explicitly specify a cost function without studying the attention allocation problem, which is the focus of this paper.

[23]Hellwig and Veldkamp (2009) adopt the same set of assumptions, but they have a different specification for the way agents reduce the observation nosies. In their work, they allow agents to pay a fixed cost to obtain a noisy signal of the fundamental, which reduces the variance of the observation noise from infinity to zero. In contrast, we allow agents to choose the variance of the observation noises by spending capacity on those signals. Furthermore, the cost function of acquiring information is also different in their work.

[24]Private information can be interpreted alternatively. In an example of an island economy, island-level productivity can be a noisy signal about the aggregate productivity. Firms can obtain a noisy signal about local productivity by paying attention to it. The island-specific component, i.e., the difference between idiosyncratic island productivity and aggregate productivity, corresponds to the sender's noise in the private signal in our model. An increase in attention to the island-level productivity only reduces the observation noise.

1.7.1. Observe the fundamental directly

We first consider a case in which agents can pay attention to observe both the fundamental directly and a noisy public signal about the fundamental. Interestingly, this scenario is a special case in our benchmark model, in which the variance of the noise contained in the private signal σ_x^2 is 0. In this case, the relative accuracy of the public signal takes an extreme value $\nabla = \rho$. (Recall that ∇ is always bounded by ρ and $1/\rho$.)

Two results arise from this extreme case. First, spending *all* of the endowed capacity on observing the fundamental directly is *always* an equilibrium and socially optimal.[25] Intuitively, to estimate the fundamental θ, spending capacity to observe θ directly is always more efficient than learning a noisy signal z about it, in terms of reducing the observation noises. At the limit, when agents process an infinite amount of capacity, they only want to spend the capacity to observe θ, in which case they can obtain the value of the fundamental. From the perspective of society, obtaining a clearer signal about the fundamental and discarding the public signal entirely is also the optimum, given that coordination is socially costly.

Second, agents may also coordinate on an inefficient equilibrium, in which everybody pays all his attention to the public signal and does not observe the fundamental at all. That case arises, only when the coordination motive is sufficiently strong, i.e., $(1 + \sqrt{1-\rho^2})/2 < \alpha < 1$ and the total capacity is not too large or too small.[26] The intuition is also not so different from the reason why multiple equilibria emerge in the benchmark case (discussed in Section 1.3.4).[27]

1.7.2. Observe two public signals

The assumption that the two information sources are of different publicity is important for our results. To highlight its effect, in this section we investigate an alternative setting where the two signals are both public but differ only in

[25] In other words, it holds that $\gamma(0,0) < 1$ for any $\hat{\kappa}$.

[26] In this case, even though ∇ takes the smallest possible value ρ, it is still possible that $\nabla_1 < \nabla = \rho < \nabla_0$. Recall that both ∇_1 and ∇_0 vary in $\hat{\kappa}$. In this case, $\hat{\nabla}$, that is, the minimum of ∇_1, is smaller than ρ. See Figure 1.A.1(e) for an illustration of the attention allocation pattern, when $\nabla = \rho$.

[27] When the strategic concern is strong, agent i is severely punished for deviating from the strategy of observing the public signal only, on the condition that everyone else adopts this strategy. To ensure that it is an equilibrium, the amount of capacity available cannot be too large; otherwise it is too costly for agent i to follow this strategy, that is, spending capacity on observing fundamental directly is more efficient in terms of learning θ. The amount of capacity cannot be too small either; otherwise the effect of relative accuracy dominates and agent i chooses to deviate.

precision.[28]

Some results from this case are the same as those from our benchmark model. First, in this case, agents focus on the relatively more precise signal when the capacity available is rather small and diversify when it is very large. Second, it is also the case that multiple equilibria may emerge, when capacity is in the intermediate range. The intuitions gained from our benchmark model can be applied to explain these two results.

However, in this setting, it is not possible that agents pay full attention to the less accurate signal and ignore the relatively more precise one, unless multiple equilibria exist. In other words, the key mechanism of attention mis-allocation in our benchmark model does not exist in this setting. Intuitively, if one of the signals is private and more precise and the other is public but less precise, it can be the case that agents may decrease their attention to, or totally ignore, the more precise private signal when the available capacity increases, because the desire of coordination dominates the effect of diminishing returns and relative accuracy. However, in the case of two public signals, that cannot be the case because both signals are public and can help coordinate agents' actions. Therefore, agents either pay full attention to the more precise public signal (when the effect of relative accuracy dominates), or divide their attention between the two public signals (when the effect of diminishing returns dominates).

1.8. Conclusion

There has been a recent surge of interest in modeling information acquisition and the endogenous information structure in macroeconomic environments. See Veldkamp (2011) for a textbook treatment on this topic and Hellwig, Kohls, and Veldkamp (2012) for an excellent review. However, fewer studies have touched on the welfare implications of information acquisition in this class of economies. This study focuses exclusively on a range of welfare issues in beauty contest models, in a context where agents are rationally inattentive and therefore optimally allocate a limited amount of attention between correlated private and public signals.

We fully characterize the sufficient and necessary conditions for the equilibrium uniqueness and multiplicity, and show that the attention allocation and the number of signals that agents decide to observe are not necessarily monotonic, in response to the increase in the capacity of processing information. Un-

[28] We can provide a full characterization of this case upon request.

like the literature, we also highlight the role of the correlation between two signals, which critically affects the equilibrium uniqueness and multiplicity, along with the welfare properties in this model. Further, we show that in this setting, when capacity increases, the social welfare of this economy may not necessarily increase. In fact, it can decrease as a result of attention misallocation. Interestingly, social welfare can be even higher when agents possess a finite amount of capacity than when they have an infinite amount of capacity.

Appendix of Chapter 1

1.A. Proofs

Proof of Lemma 1.1. First, we solve for the weighting rule adopted by all of the other agents, on the condition that their attention allocation is $(\hat{\kappa}_x^*, \hat{\kappa}_z^*)$:

$$\Pi_x^* = \frac{\phi_x^*}{\phi_x^* + \phi_z^* + \phi_\theta^*}, \qquad \Pi_z^* = \frac{\phi_z^*}{\phi_x^* + \phi_z^* + \phi_\theta^*}, \qquad \Pi_\theta^* = \frac{\phi_\theta^*}{\phi_x^* + \phi_z^* + \phi_\theta^*},$$

where

$$\phi_x^* = \frac{1}{\sigma_x^2 + (\omega_x^*)^2}, \qquad \phi_z^* = \frac{1}{(1-\alpha)\,\sigma_z^2 + (\omega_z^*)^2}, \qquad \phi_\theta^* = \frac{1}{(1-\alpha)\,\sigma^2},$$

and

$$\omega_x^* = \sqrt{\frac{(\sigma_x^2 + \sigma^2)\,(1-\rho^2)}{\exp(2\hat{\kappa}_x^*) - 1}}, \qquad \omega_z^* = \sqrt{\frac{(\sigma_z^2 + \sigma^2)\,(1-\rho^2)}{\exp(2\hat{\kappa}_z^*) - 1}}.$$

Second, we solve for the optimal action rule for agent i, i.e., $(\Pi_{z,i}^*, \Pi_{x,i}^*, \Pi_{\theta,i}^*)$, conditional on the others' allocation strategy $(\hat{\kappa}_x^*, \hat{\kappa}_z^*)$ and his own $(\hat{\kappa}_{x,i}, \hat{\kappa}_{z,i})$. It is the solution to the following optimization problem,

$$\max_{\Pi_{x,i}, \Pi_{z,i}, \Pi_{\theta,i}} E[u_i] \quad s.t. \left(\hat{\kappa}_{z,j}, \hat{\kappa}_{x,j}\right) = (\hat{\kappa}_z^*, \hat{\kappa}_x^*) \text{ for all } j \neq i,$$

where $E[u_i]$ is given by equation (1.14). First order conditions imply that

$$\Pi_{x,i}^* = \Pi_x^* + \frac{(c_1\Pi_\theta^* - c_2\Pi_x^*)\,(c_1 + c_5 + c_3 + c_4) - (c_1\Pi_\theta^* - c_3\Pi_z^*)\,(c_1 + c_5)}{(c_1 + c_5 + c_2)\,(c_1 + c_5 + c_3 + c_4) - (c_1 + c_5)^2}, \tag{1.23}$$

$$\Pi_{z,i}^* = \Pi_z^* + \frac{(c_1\Pi_\theta^* - c_3\Pi_z^*)\,(c_1 + c_5 + c_2) - (c_1\Pi_\theta^* - c_2\Pi_x^*)\,(c_1 + c_5)}{(c_1 + c_5 + c_2)\,(c_1 + c_5 + c_3 + c_4) - (c_1 + c_5)^2}. \tag{1.24}$$

where

$$c_1 = (1-\alpha)\,\sigma^2, \qquad c_2 = \sigma_x^2 + \omega_{x,i}^2, \qquad c_3 = (1-\alpha)\,\sigma_z^2 + \omega_{z,i}^2,$$
$$c_4 = \alpha\sigma_z^2, \qquad c_5 = \alpha\sigma^2.$$

Therefore, the relative marginal return of attention on the public signal γ can be re-written by replacing $\Pi_{x,i}^*$ and $\Pi_{z,i}^*$ with (1.23) and (1.24).

$$\gamma = \left[\frac{C_1 \exp\left(2\hat{\kappa}_{z,i}\right) + C_2 \exp\left(2\hat{\kappa}\right)}{C_3 \exp\left(2\hat{\kappa}_{z,i}\right) + C_4}\right]^2 \frac{1}{\nabla^2 \exp\left(2\hat{\kappa}\right)}.$$

where

$$\begin{aligned}
C_1 &= \left[(1-\alpha)\,\Pi_\theta^* + (1-\alpha)\,\Pi_z^* + \Pi_x^*\right](1-\nabla\rho)\,,\\
C_2 &= \alpha\left(\frac{1}{\rho^2}-1\right)\Pi_z^* + \left[(1-\alpha)\,\Pi_\theta^* + (1-\alpha)\,\Pi_z^* + \Pi_x^*\right]\left(\frac{\nabla}{\rho}-1\right),\\
C_3 &= \left[(1-\alpha)\,\Pi_\theta^* + (1-\alpha)\,\Pi_z^* + \Pi_x^*\right]\left(\frac{1}{\nabla\rho}-1\right),\\
C_4 &= \alpha\left(\frac{1}{\nabla\rho}-\frac{\rho}{\nabla}\right)\Pi_z^* + \left[(1-\alpha)\,\Pi_\theta^* + (1-\alpha)\,\Pi_z^* + \Pi_x^*\right]\left(1-\frac{\rho}{\nabla}\right).
\end{aligned}$$

Therefore, we can show

$$\frac{\partial\gamma\left(\hat{\kappa}_{z,i},\hat{\kappa}_z^*\right)}{\partial\hat{\kappa}_{z,i}} < 0, \qquad \frac{\partial\gamma\left(\hat{\kappa}_{z,i},\hat{\kappa}_z^*\right)}{\partial\nabla} > 0, \qquad \frac{\partial\gamma\left(\hat{\kappa}_{z,i},\hat{\kappa}_z^*\right)}{\partial\alpha} > 0.$$

■

Proof of Proposition 1.1. The first part of Lemma 1.1 implies that the best response of agent i to the allocation strategy adopted by others is unique. Therefore, the allocation $(\hat{\kappa}_x,\hat{\kappa}_z) = (\hat{\kappa},0)$ constitutes a symmetric equilibrium if and only if $\gamma(0,0) < 1$. That is,

$$\nabla \leq \frac{\exp\left(\hat{\kappa}\right)\rho + 1}{\exp\left(\hat{\kappa}\right) + \rho} \equiv \nabla_0.$$

Similarly, the allocation $(\hat{\kappa}_x,\hat{\kappa}_z) = (0,\hat{\kappa})$ constitutes a symmetric equilibrium if and only if $\gamma(\hat{\kappa},\hat{\kappa}) > 1$. That is,

$$\nabla \geq \frac{(1-\alpha)\left(\exp(2\hat{\kappa})-1\right) + \left(1-\rho^2\right)}{(1-\alpha)\left(\exp(2\hat{\kappa})-1\right)\rho + \exp(\hat{\kappa})\left(1-\rho^2\right)} \equiv \nabla_1.$$

A symmetric equilibrium with diversification must be such that $\hat{\kappa}_z^* \in (0,\hat{\kappa})$ and it exists if and only if $\gamma\left(\hat{\kappa}_z^*,\hat{\kappa}_z^*\right) = 1$ where we have

$$\gamma\left(\hat{\kappa}_z^*,\hat{\kappa}_z^*\right) = \frac{\left(\left(\frac{\nabla}{\rho}-1\right)\left(\exp(2\hat{\kappa})-\exp(2\hat{\kappa}_z^*)\right) + \nabla\left(\frac{1}{\rho}-\rho\right)\exp(2\hat{\kappa}_z^*)\right)^2}{\exp(2\hat{\kappa})\left((1-\alpha)\left(\frac{1}{\rho}-\nabla\right)\left(\exp(2\hat{\kappa}_z^*)-1\right) + \left(\frac{1}{\rho}-\rho\right)\right)^2}. \tag{1.25}$$

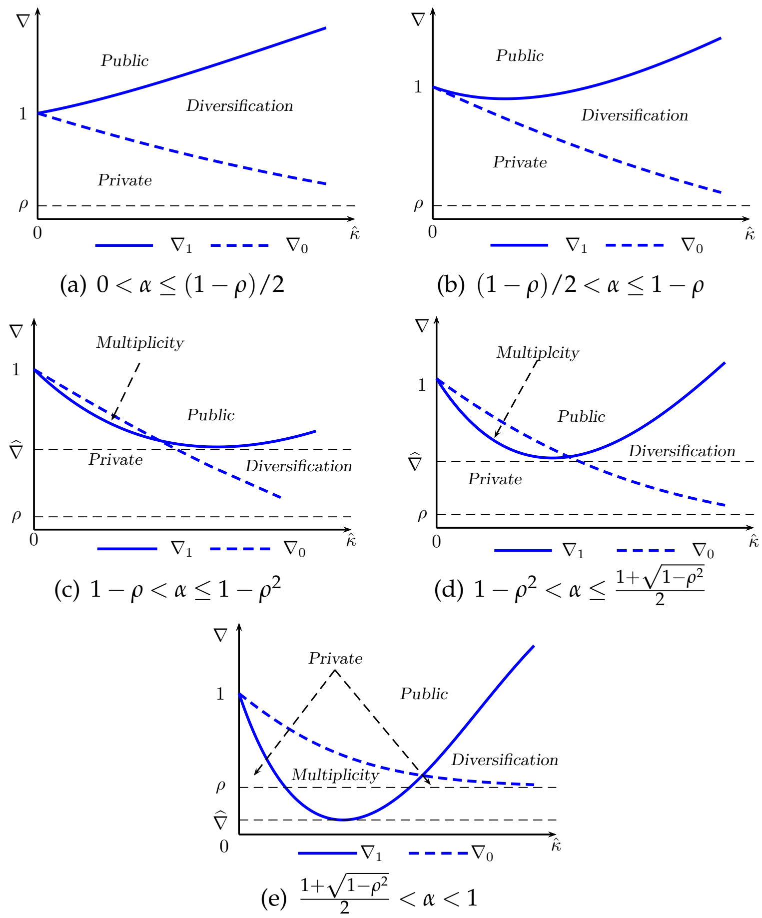


Figure 1.A.1. *Patterns of bounds* ∇_0 *and* ∇_1.

Such an equilibrium arises if

$$\nabla \in (\min\{\nabla_0, \nabla_1\}, \max\{\nabla_0, \nabla_1\}).$$

The optimal allocation is given by (1.18). Obviously, the equilibrium must be unique, if $\nabla_0 < \nabla_1$, which also implies $\alpha < 1 - \frac{\rho}{\exp(\hat{\kappa})}$. In other words, multiple equilibria emerge if and only if the condition (1.20) holds. ■

Proof of Proposition 1.2. When $\hat{\kappa}$ is sufficiently large, ∇_1 monotonically increases and $\lim_{\hat{\kappa}\to+\infty} \nabla_1 = \frac{1}{\rho}$ while ∇_0 monotonically decreases and $\lim_{\hat{\kappa}\to+\infty} \nabla_0 = \rho$. Therefore, for any ∇, when $\hat{\kappa}$ is sufficiently large, it holds that $\nabla \in (\nabla_0, \nabla_1)$. According to Proposition 1.1, the equilibrium is unique and $0 < \hat{\kappa}_z^* < \hat{\kappa}$. Further, the last part of this proposition can be obtained from equation (1.18). ■

Proof of Proposition 1.3. The complete characterization of the equilibrium attention allocation can be summarized in the following claims.

Claim 1: If the relative accuracy is sufficiently low (i.e., $\nabla < \hat{\nabla}$), agents specialize in learning the private signal and then eventually diversify their attention when the capacity increases, where

$$\hat{\nabla} = \frac{1}{\rho} - \frac{1}{\rho} \frac{1}{2\sqrt{\frac{\alpha(1-\alpha)}{(1-\rho^2)}}\rho + \frac{2(1-\alpha)}{(1-\rho^2)}\rho^2 + 1};$$

$\hat{\kappa}_z^* / \hat{\kappa}$ can be either monotonically increasing or hump-shaped.

Proof. We first establish some properties of ∇_0 and ∇_1. As illustrated in Figure 1.A.1, there are in total five patterns, according to the combinations of α and ρ. For any $\hat{\kappa} > 0$, the bounds ∇_0 and ∇_1 can be characterized as follows,

1. $\lim_{\hat{\kappa} \to +\infty} \nabla_0 = \rho$ and $\lim_{\hat{\kappa} \to +\infty} \nabla_1 = 1/\rho$.
2. $\nabla_0(0) = \nabla_1(0) = 1$.
3. ∇_0 monotonically decreases in $\hat{\kappa}$.
4. ∇_1 may or may not be monotone:
   (i) If $0 < \alpha < \frac{1-\rho}{2}$, ∇_1 monotonically increases in $\hat{\kappa}$. Otherwise, ∇_1 decreases and then increases, reaching the trough at $\hat{\kappa} = \hat{\hat{\kappa}}$, where
   $$\hat{\hat{\kappa}} = \ln\left(\rho + \sqrt{\alpha\,(1-\rho^2)\,/\,(1-\alpha)}\right).$$
   (ii) If $\frac{1-\rho}{2} < \alpha < 1 - \rho$, ∇_1 is always larger than ∇_0 for any $\hat{\kappa}$.
   (iii) If $1 - \rho < \alpha$, ∇_0 and ∇_1 cross only once at $\hat{\kappa} = \ln\left(\frac{\rho}{1-\alpha}\right)$, on the condition that $\hat{\kappa}$ is positive. Further, ∇_1 is smaller than ∇_0 if and only if $\hat{\kappa} < \ln(\frac{\rho}{1-\alpha})$.
   (iv) If $1 - \rho < \alpha < 1 - \rho^2$, ∇_0 and ∇_1 cross on the left side of $\hat{\hat{\kappa}}$; if $1 - \rho^2 < \alpha < 1$, they cross on the right side of $\hat{\hat{\kappa}}$.

We can show the first three items by using the expression in equation (1.19). The last item can be verified by noting that,

$$\frac{\mathrm{d}\nabla_1}{\mathrm{d}\hat{\kappa}} \quad \propto \quad (1-\alpha)\exp(2\hat{\kappa}) - 2\,(1-\alpha)\,\rho\exp(\hat{\kappa}) + \rho^2 - \alpha.$$

Denote $\hat{\nabla} \equiv \nabla_1\left(\hat{\kappa}\right)$ and $\tilde{\nabla} \equiv \nabla_1\left(\ln\left(\frac{\rho}{1-\alpha}\right)\right)$, we obtain

$$\hat{\nabla} = \frac{1}{\rho} - \frac{1}{\rho}\frac{1}{2\sqrt{\frac{\alpha(1-\alpha)}{(1-\rho^2)}}\rho + \frac{2(1-\alpha)}{(1-\rho^2)}\rho^2 + 1}; \qquad \tilde{\nabla} = \frac{1}{\rho} - \frac{1}{\rho}\frac{\alpha\left(1-\rho^2\right)}{1-(1-\alpha)^2}.$$

Intuitively, $\hat{\nabla}$ is the minimum of ∇_1 and $\tilde{\nabla}$ is such that $\tilde{\nabla} = \nabla_1\left(\ln\left(\frac{\rho}{1-\alpha}\right)\right) = \nabla_0\left(\ln\left(\frac{\rho}{1-\alpha}\right)\right)$. If $\nabla \in (\rho, \hat{\nabla})$, it holds that $\nabla < \nabla_1$ for any $\hat{\kappa}$; and there exists a cutoff $\hat{\kappa}_0$, such that for any $\hat{\kappa} \in (0, \hat{\kappa}_0)$, $\nabla < \nabla_0$ and for any $\hat{\kappa} \geq \hat{\kappa}_0$, $\nabla \geq \nabla_0$. According to Proposition 1.1, the first part of this Claim is shown.

Regarding the pattern of attention allocation $\hat{\kappa}_z^*/\hat{\kappa}$, it can be categorized in the following two cases:

1. When the coordination motive is not so high, i.e., $0 < \alpha < 1-\rho$, $\hat{\kappa}_z^*/\hat{\kappa}$ is monotonically increasing in $\hat{\kappa}$, if $\nabla < \overline{\nabla}$, where $\overline{\nabla} = [\rho + (1-\alpha)]/[(1-\alpha)\rho + 1]$; and is hump-shaped in $\hat{\kappa}$, if $\overline{\nabla} < \nabla < 1$.

2. When the coordination motive is high, i.e., $1-\rho < \alpha$, $\hat{\kappa}_z^*/\hat{\kappa}$ is hump-shaped in $\hat{\kappa}$, if $\overline{\nabla} < \nabla < \min\{\tilde{\nabla}, \hat{\nabla}\}$; and is monotonically increasing in $\hat{\kappa}$, if $\nabla < \overline{\nabla}$.

The details of the proof of the above two cases are contained in the Technical Appendix. ∎

Claim 2: Suppose the coordination motive is strong, i.e., $(1-\rho)/2 < \alpha \leq 1-\rho$, and the relative accuracy is not extremely high, i.e., $\hat{\nabla} < \nabla < 1$. Agents re-allocate their attention in the following fashion. When capacity is low, they specialize in learning the private signal only, then diversify their attention allocation and then specialize in learning the public signal only before eventually diversifying again.

Proof. The proof is similar to that of Claim 1. See Figure 1.A.1(b) for illustration. ∎

Claim 3: Suppose the coordination motive is very strong, i.e., $1-\rho < \alpha < 1$, and the relative accuracy is not extremely high, i.e., $\tilde{\nabla} \leq \nabla < 1$. Agents re-allocate their attention in the following fashion. When the capacity is sufficiently low, they focus only on the private signal. When the capacity is higher, they may coordinate on one of the three equilibria. When there is a further increase in capacity, they pay attention only to the public signal. And when the capacity is sufficiently high, they eventually diversify.

Proof. We show that $\tilde{\nabla} < 1$ if and only if $\rho > 1 - \alpha$. The proof is similar to that of Claim 1. See Figure 1.A.1(c), (d) and (e) for illustration. ∎

Claim 4: Suppose the coordination motive is very strong, i.e., $1 - \rho < \alpha \leq 1$, and the relative accuracy is low, i.e., $\max\{\rho, \hat{\nabla}\} \leq \nabla < \tilde{\nabla}$. If $1 - \rho < \alpha \leq 1 - \rho^2$, they allocate their attention in the same fashion as that in Claim 2. If $1 - \rho^2 < \alpha \leq 1$, when capacity is low, agents specialize in learning the private signal only. When the capacity is higher, they may coordinate on one of the three equilibria. When there is a further increase in capacity, they pay attention only to the private signal again. And when the capacity is sufficiently high, they eventually diversify. Note that $\rho < \hat{\nabla}$, if and only if $\alpha < (1 + \sqrt{1 - \rho^2})/2$.

Proof. The proof is similar to that of Claim 1. See Figure 1.A.1(d) and (e) for illustration. ∎

∎

Proof of Lemma 1.2. This proof offers a sufficient condition under which Lemma 1.2 holds. We consider the following two cases: (i) $2\alpha + \rho > 1$, $(1 - \alpha)\exp(\hat{\kappa}) > \rho$ and $\nabla < 1$; (ii) $\alpha + \rho > 1$ and $\nabla < \hat{\nabla}$. Under these two cases, if $\hat{\kappa} > \hat{\kappa}_0$, $0 \leq \hat{\kappa}_x^* < \hat{\kappa}$. Therefore, we can show,

$$\begin{aligned}
\frac{\mathrm{d}\hat{\kappa}_x^*}{\mathrm{d}\hat{\kappa}} &\propto (\nabla - \rho)(1 - \alpha)\exp(2\hat{\kappa}) - \left[\left(1 - \rho^2\right) - (1 - \rho\nabla)(1 - \alpha)\right](1 - \alpha)\exp(\hat{\kappa}) \\
&- \left((1 - \alpha)\exp(\hat{\kappa}) - \rho\right)\left[\left(1 - \rho^2\right) - (1 - \alpha)(1 - \nabla\rho)\right] \\
&< \left[(\nabla - \rho)\exp(\hat{\kappa}) - \left(1 - \rho^2\right) + (1 - \rho\nabla)(1 - \alpha)\right](1 - \alpha)\exp(\hat{\kappa}).
\end{aligned}$$

Therefore, $\mathrm{d}\hat{\kappa}_x^*/\mathrm{d}\hat{\kappa} < 0$, if $\left[(\nabla - \rho)\exp(\hat{\kappa}) - \left(1 - \rho^2\right) + (1 - \rho\nabla)(1 - \alpha)\right] < 0$. This holds true, if

$$\frac{1 - \nabla\rho}{\nabla - \rho} < \exp(\hat{\kappa}) < \frac{\left(1 - \rho^2\right) - (1 - \rho\nabla)(1 - \alpha)}{\nabla - \rho}.$$

The first inequality must hold so that $\hat{\kappa} > \hat{\kappa}_0$. The second inequality can hold on the condition that α is sufficiently large; that is,

$$\alpha > 1 - \frac{(1-\rho^2)}{(1-\rho\nabla)}.$$

■

Proof of Proposition 1.4. A simple calculation leads to

$$\frac{\mathrm{d}\nabla_1}{\mathrm{d}\rho} = \frac{(\exp(2\hat{\kappa})-1)(1-\alpha)\left[-(1-\alpha)(\exp(2\hat{\kappa})-1)+2\rho\exp(\hat{\kappa})-\left(1+\rho^2\right)\right]}{\left[(1-\alpha)(\exp(2\hat{\kappa})-1)\rho+\exp(\hat{\kappa})(1-\rho^2)\right]^2}.$$

Let $T_1 = -(1-\alpha)(\exp(2\hat{\kappa})-1)+2\rho\exp(\hat{\kappa})-\left(1+\rho^2\right)$. If and only if $\rho < (1-\alpha)\exp(\hat{\kappa})$, T_1 decreases in $\hat{\kappa}$ and $T_1 < 0$. In other words, $\mathrm{d}\nabla_1/\mathrm{d}\rho < 0$ for any $\hat{\kappa} > \ln(\rho/(1-\alpha))$. This implies that $\hat{\kappa}_1$ increases in ρ. Similarly, we can show that ∇_0 is an increasing function of ρ, and therefore $\hat{\kappa}_0$ increases in ρ.

Let $l(\hat{\kappa})$ be the difference between the expected utility of adopting the strategy $\hat{\kappa}_z^* = \hat{\kappa}$ and that of $\hat{\kappa}_z^* = 0$, when condition (1.20) holds. The cutoff $\hat{\kappa}_s$ is such that $l(\hat{\kappa}_s) = 0$. It implies that $\hat{\kappa}_z^* = \hat{\kappa}$ if and only if

$$l(\hat{\kappa}) = \left[1+\frac{(1-\rho^2)}{\exp(2\hat{\kappa})-1}\right]\nabla^2 - \alpha\rho\nabla - \left[\frac{(1-\rho^2)}{\exp(2\hat{\kappa})-1}+(1-\alpha)\right] > 0.$$

Under this circumstance, it is straightforward to show the following facts: $l(\hat{\kappa})$ is strictly increasing in $\hat{\kappa}$, $\lim_{\hat{\kappa}\to 0} l(\hat{\kappa}) < 0$ and $\lim_{\hat{\kappa}\to+\infty} l(\hat{\kappa}) > 0$. Therefore, there is a unique $\hat{\kappa}_s > 0$, such that $l(\hat{\kappa}_s) = 0$, where

$$\hat{\kappa}_s = \ln\left(\sqrt{\frac{(1-\rho^2)(1-\nabla^2)}{\alpha(1-\rho\nabla)-(1-\nabla^2)}+1}\right).$$

By noting that $\hat{\kappa}_s$ decreases in ρ, this proposition is shown. ■

Proof of Proposition 1.5. We show that social welfare can be decreasing in the case where $\alpha + \rho > 1$ and $\tilde{\nabla} < \nabla < 1$. If $\hat{\kappa}_z^* = 0$ or $\hat{\kappa}_z^* = \hat{\kappa}$, social welfare is calculated by the following

$$E\left[\mathcal{W}^s\right] = -\left[\frac{1+(1-\alpha)\frac{\phi_x}{\phi_\theta}}{\left(\frac{\phi_x}{\phi_\theta}+1\right)^2}\right]\sigma^2,$$

$$E\left[\mathcal{W}^s\right] = -\left[\frac{1+(1-\alpha)\frac{\phi_z}{\phi_\theta}}{\left(\frac{\phi_z}{\phi_\theta}+1\right)^2}\right]\sigma^2 - \alpha\frac{\left(\frac{\phi_z}{\phi_\theta}\right)^2}{\left(\frac{\phi_z}{\phi_\theta}+1\right)^2}\sigma_z^2,$$

when $\hat{\kappa} < \hat{\kappa}_s$, $\hat{\kappa}_z^* = 0$ and $\mathrm{d}E\left[\mathcal{W}\left(\mathbf{a},\theta\right)\right]/\mathrm{d}\hat{\kappa} > 0$. To see this, we notice that ϕ_x increases in $\hat{\kappa}$ and $E\left[\mathcal{W}^s\right]$ increases in ϕ_x,

$$\frac{\mathrm{d}E\left[\mathcal{W}^s\right]}{\mathrm{d}\phi_x} = \frac{\sigma^2\left[(1+\alpha)+(1-\alpha)\frac{\phi_x}{\phi_\theta}\right]}{\left(1+\frac{\phi_x}{\phi_\theta}\right)^3\phi_\theta} > 0.$$

Similarly, when $\hat{\kappa}_s < \hat{\kappa} < \hat{\kappa}_1$, $\hat{\kappa}_z^* = \hat{\kappa}$ and social welfare increases in $\hat{\kappa}$. When $\hat{\kappa} = \hat{\kappa}_s$, agents are indifferent of specialization in private or public signals, which implies that $\phi_x = \phi_z$, and social welfare discontinuously decreases at $\hat{\kappa} = \hat{\kappa}_s$. Because $E\left[\mathcal{W}^s\right]$ monotonically increases in $\hat{\kappa}$, when $\hat{\kappa} < \hat{\kappa}_s$, there must exist $\hat{\kappa}_a$ and $\hat{\kappa}_b$ such that $E\left[\mathcal{W}^s(\hat{\kappa}_a)\right] > E\left[\mathcal{W}^s(\hat{\kappa}_b)\right]$ and $\hat{\kappa}_a < \hat{\kappa}_s < \hat{\kappa}_b$. ∎

Proof of Proposition 1.6. Except in the case where $\alpha+\rho > 1$ and $\tilde{\nabla} < \nabla < 1$, we can show that $\hat{\kappa}_x^* = \hat{\kappa}_0$ and $\hat{\kappa}_z^* = 0$, when $\hat{\kappa} = \hat{\kappa}_0$. To show $E\left[\mathcal{W}^s(\hat{\kappa}_0)\right] > E\left[\mathcal{W}^{ms}\right]$, we only need to show $f \equiv E\left[\mathcal{W}^s(\hat{\kappa}_0)\right] - E\left[\mathcal{W}^{ms}\right] > 0$, where

$$\begin{aligned} f = & \frac{1}{\left(\frac{\rho}{(\nabla-\rho)}+\frac{1}{(1-\alpha)(1-\nabla\rho)}\right)} - \alpha\frac{\frac{\rho}{(\nabla-\rho)}}{\left(\frac{\rho}{(\nabla-\rho)}+\frac{1}{(1-\alpha)(1-\nabla\rho)}\right)^2} \\ & - \frac{1}{\left(\frac{\rho(1-\nabla^2)}{(1-\rho\nabla)(\nabla-\rho)}+\frac{1}{1-\alpha}\right)} - \alpha\frac{\frac{\rho\left(1-\nabla^2\right)}{(1-\rho\nabla)(\nabla-\rho)}}{\left(\frac{\rho(1-\nabla^2)}{(1-\rho\nabla)(\nabla-\rho)}+\frac{1}{1-\alpha}\right)^2}. \end{aligned}$$

To provide a sufficient condition under which the inequality holds, we denote

$$f(\alpha) = f_1(\alpha) + f_2(\alpha),$$

where

$$f_1(\alpha) = \left(\frac{1}{1-\rho\nabla} - 1\right)(1-\alpha)^2\left(-\alpha^2+4\alpha-2\right)\left(\frac{\rho}{\nabla-\rho}\right)^2$$
$$f_2(\alpha) = \frac{(1-\alpha)^2\rho\nabla}{1-\rho\nabla}\left[\left(\frac{1}{1-\rho\nabla}-1\right)(\alpha-2)^2+2\left(\frac{1}{1-\rho\nabla}-1\right)^2-2-\frac{1}{1-\rho\nabla}\frac{(2-\alpha)^2}{(1-\alpha)^2}\right]$$

$f_1(\alpha) > 0$ if and only if $-\alpha^2+4\alpha-2>0$, or $\alpha > 2-\sqrt{2}$. When ∇ is sufficiently low and close to ρ, $\frac{\rho}{\nabla-\rho}$ can be arbitrarily large and $\frac{1}{1-\rho\nabla}$ is close to a constant $\frac{1}{1-\rho^2}$. Therefore, $f_1(\alpha)$ can be arbitrarily large and $f_2(\alpha)$ is close to a constant. Moreover, it must hold that $\rho < (1-\alpha)\exp(\hat{\kappa}_0)$ or,

$$\frac{(1-\alpha)}{\rho} \quad > \quad \frac{\nabla-\rho}{1-\rho\nabla}.$$

This holds when ∇ is low enough. ∎

1.B. Technical Appendix

Proof of Proposition 1.3, Part 2 of Claim 1. The following lemma is particularly useful for our characterization.

Lemma 1.3. When the equilibrium is unique, the fraction of attention paid to the public signal, i.e., $\hat{\kappa}_z^*/\hat{\kappa}$, strictly increases in $\hat{\kappa}$ if and only if $0 < \hat{\kappa}_z^*/\hat{\kappa} < 1$ and

$$\frac{\hat{\kappa}_z^*}{\hat{\kappa}} - F(\hat{\kappa}) < 0, \tag{1.26}$$

where

$$F \equiv \frac{1}{2} + \frac{1}{2}\frac{T}{(1-\alpha)(1-\rho\nabla)(\nabla-\rho)} \cdot \frac{1}{1-\frac{(1-\rho^2)-(1-\alpha)(1-\rho\nabla)}{(\nabla-\rho)\exp(\hat{\kappa})}} \cdot \frac{1}{\exp(\hat{\kappa})-\frac{\rho}{1-\alpha}},$$
$$T = (1-\rho\nabla)\,\alpha\left[(1-\rho\nabla)(1-\alpha)-\rho(\nabla-\rho)\right].$$

Proof. We re-write equation (1.18) as follows,

$$\frac{\hat{\kappa}_z^*}{\hat{\kappa}} = \frac{1}{2} + \frac{\frac{1}{2}\ln\left[\frac{(1-\alpha)(1-\rho\nabla)+(\nabla-\rho)\exp(\hat{\kappa})-\left(1-\rho^2\right)}{(1-\alpha)(1-\rho\nabla)\exp(\hat{\kappa})+(\nabla-\rho)-\nabla(1-\rho^2)}\right]}{\ln\left[\exp(\hat{\kappa})\right]}. \tag{1.27}$$

Derive its derivative with respect to $\exp(\hat{\kappa})$ and we find that $\partial(\hat{\kappa}_z^*/\hat{\kappa})/\partial\exp(\hat{\kappa}) > 0$ if and only if the condition (1.26) holds. ■

We first show that in the case of $\nabla > 1$, $\hat{\kappa}_z^*/\hat{\kappa}$ monotonically decreases in $\hat{\kappa}$, when $\hat{\kappa} > \hat{\kappa}_1$. There are three sub-cases.

Case 1: $\alpha < 1-\rho$ and $\nabla \in \left(\frac{1-\alpha+\rho^2}{(1-\alpha)\rho+\rho}, \frac{1}{\rho}\right)$. We can verify that $T < 0$ and further

$$0 > \ln\left(\frac{\rho}{1-\alpha}\right) > \ln\left(\frac{(1-\rho^2)-(1-\alpha)(1-\rho\nabla)}{\nabla-\rho}\right),$$

which implies that $F(\hat{\kappa})$ is strictly increasing and approaches $1/2$ from below when $\hat{\kappa}$ approaches $+\infty$.

We can also show that $\hat{\kappa}_z^*/\hat{\kappa}$ approaches $1/2$ from above, because the second term in (1.27) is positive when $\hat{\kappa}$ approaches $+\infty$, i.e.,

$$\begin{aligned} &\lim_{\hat{\kappa}\to+\infty} \ln\left[\frac{(1-\alpha)(1-\rho\nabla)+(\nabla-\rho)\exp(\hat{\kappa})-(1-\rho^2)}{(1-\alpha)(1-\rho\nabla)\exp(\hat{\kappa})+(\nabla-\rho)-\nabla(1-\rho^2)}\right] \\ = \;& \ln\left[\frac{\nabla-\rho}{(1-\alpha)(1-\rho\nabla)}\right] > 0. \end{aligned}$$

Further, the first part of Proposition 1.3 implies that $\lim_{\hat{\kappa}\to\hat{\kappa}_1^+}\hat{\kappa}_z^*/\hat{\kappa} = 1$. We can show that $\hat{\kappa}_z^*/\hat{\kappa} - F(\hat{\kappa}) > 0$, for any $\hat{\kappa} > \hat{\kappa}_1$, by constructing a contradiction. Suppose there exists $\hat{\kappa}'$ such that $\hat{\kappa}_z^*(\hat{\kappa}')/\hat{\kappa}' < F(\hat{\kappa}')$. Lemma 1.3 implies that it must hold that $\hat{\kappa}_z^*/\hat{\kappa}$ approaches $1/2$ from below. A contradiction. This fact further implies that $\hat{\kappa}_z^*/\hat{\kappa}$ monotonically decreases, by using Lemma 1.3 again.

Case 2: $\alpha < 1-\rho$ and $\nabla \in \left(1, \frac{1-\alpha+\rho^2}{(1-\alpha)\rho+\rho}\right)$. In this case, we can show that $F(\hat{\kappa})$ is strictly decreasing and approaches $1/2$ from above, when $\hat{\kappa}$ approaches $+\infty$. Further, it must hold that $F(\hat{\kappa}_1) \leq 1$, that is because $\hat{\kappa}_z^*/\hat{\kappa}$ must decrease, when $\hat{\kappa}$ is slightly higher than $\hat{\kappa}_1$, according to the first part of Proposition 1.3.

Similar to the previous case, Lemma 1.3 implies that $\hat{\kappa}_z^*/\hat{\kappa} - F(\hat{\kappa})$ cannot cross zero from above; that is, $\hat{\kappa}_z^*/\hat{\kappa} - F(\hat{\kappa}) > 0$ for any $\hat{\kappa} > \hat{\kappa}_1$. In other words, $\hat{\kappa}_z^*/\hat{\kappa}$ decreases monotonically.

Case 3: The proof for the case where $\alpha > 1-\rho$ and $\nabla > 1$ is similar.

We then establish that $\hat{\kappa}_z^*/\hat{\kappa}$ can be either monotonically increasing or hump-shaped when $\nabla < 1$.

Case 1: $\alpha < 1-\rho$ and $\nabla \in (\rho, \overline{\nabla})$. In this case, we can show that $F(\hat{\kappa})$ is monotonically decreasing and approaches $1/2$ from above. The first part of Proposition 1.3 implies that $\lim_{\hat{\kappa}\to\hat{\kappa}_0^+} \hat{\kappa}_z^*/\hat{\kappa} = 0$. Under this case, $\hat{\kappa}_z^*/\hat{\kappa}$ approaches $1/2$ from below because the second term in (1.27) is negative when $\hat{\kappa}$ approaches $+\infty$, i.e.,

$$\begin{aligned}
&\lim_{\hat{\kappa}\to+\infty} \ln\left[\frac{(1-\alpha)(1-\rho\nabla)+(\nabla-\rho)\exp(\hat{\kappa})-\left(1-\rho^2\right)}{(1-\alpha)(1-\rho\nabla)\exp(\hat{\kappa})+(\nabla-\rho)-\nabla\left(1-\rho^2\right)}\right] \\
=\;& \ln\left[\frac{\nabla-\rho}{(1-\alpha)(1-\rho\nabla)}\right] < 0.
\end{aligned}$$

Using similar arguments in previous cases, we can show that $\hat{\kappa}_z^*/\hat{\kappa}$ monotonically increases in $\hat{\kappa}$.

Case 2: $\alpha \in (0,1-\rho)$ and $\nabla \in (\overline{\nabla},1)$. This case differs from the previous one in that $\hat{\kappa}_z^*/\hat{\kappa}$ approaches $1/2$ from above when $\hat{\kappa} \to +\infty$. Therefore, Lemma 1.3 implies that $\hat{\kappa}_z^*/\hat{\kappa}$ must be increasing and then decreasing, i.e., it is hump-shaped.

The proofs of the remaining cases are also similar. ∎

Chapter 2

Ambiguity, Pessimism and Economic Fluctuations[1]

[1]I thank Fabrizio Zilibotti, George-Marios Angeletos, Massimo Marinacci and Nir Jaimovich for their useful discussions and comments. I would also express my gratitude for George-Marios Angeletos for the host in MIT from Feb. 2017 to Jun. 2017, when large parts of the project are conducted.

2.1. Introduction

Recessions are times with depressed market confidence and heightened divergence in belief. Before the onset of any economic crisis, markets are quite confident in the prospects of the economy and less diverged in their beliefs. When the crisis steps in, market confidence plummets and belief divergence soars. In the data, we proxy market confidence by sentiment index in Michigan Survey of Consumers and measure belief divergence by the cross-sectional dispersion of real GDP forecasts in Survey of Professional Forecasters. Depicted in Figure 2.1, over the last thirty years, all three recessions experienced by the US economy feature large swings in both market confidence and belief divergence. Natural questions of interests arise. Why the aggregate economy co-moves with market confidence and belief divergence? What explains the strong negative correlation between market confidence and belief divergence?

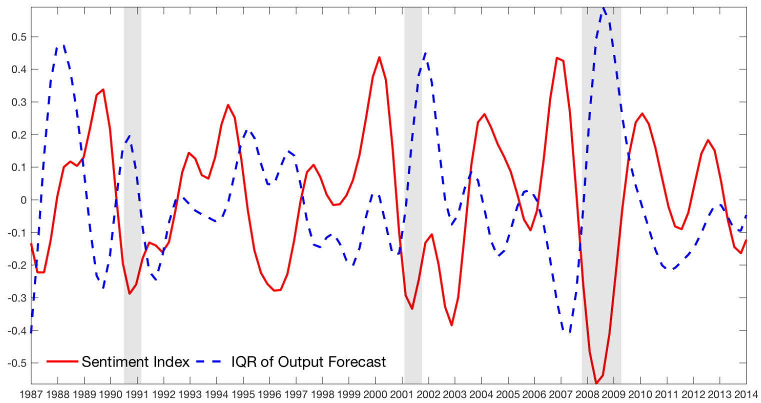


Figure 2.1. *Market Confidence and Belief Divergence.*

Note: The figure plots consumer sentiment index (red line) and cross-sectional dispersion of real GDP forecasts by professional forecasters (dashed blue line) over the period 1987Q1-2014Q4. Correlation is around -0.45 over the entire period. Both of the time series are bandpass-filtered at frequencies 6-32 quarters and re-scaled. Consumer sentiment index is from Michigan Survey of Consumers, and real GDP forecast data is from Survey of Professional Forecasters provided by the Federal Reserve Bank of Philadelphia.

Recent financial crisis highlights the importance of ambiguity[2] in driving fluctuations in asset prices and aggregate economy. Along the onset of the crisis, a sequence of events, such as the BNP Paribas announcement, the Bear Stearns rescue, the bankruptcy of Lehman Brothers, etc., made agents start to

[2]Ambiguity refers to subjective uncertainty over probabilities due to lack of ex-ante information to pin down a specific model for the economy in the course of decision making.

doubt their "ways of thinking" over the economy. Hit by these unusual events, agents felt increasingly hard to pin down a probability model for the evaluation of the outlooks of the economy. In response, investors began to question their asset pricing models, and firms began to suspect the models that were used to estimate market demand for their products. To self-insure against these doubts over models or increased ambiguity in more general, agents started to behave pessimistically, either by fire-selling or by shrinking production. These manifested as a drop in market confidence. On the other hand, these doubts over models also made it increasingly difficult to coordinate each others' beliefs. Inability to formalize a common model for the evaluation of the outlooks of the economy raised up belief divergence. In a pure narrative sense, ambiguity tends to be an important driving force in the background of co-movements across market confidence, belief divergence, and aggregate economy.

To materialize this idea, we extend the standard theories with **ambiguity averse** agents[3], who are hit by shocks that fluctuate the amount of ambiguity they perceived. The key feature of the paper is that ambiguity averse preferences are mathematically represented by the smooth model of ambiguity axiomatized by Klibanoff, Marinacci, and Mukerji (2005, 2009), within which the amount of ambiguity can be measured by the variance of the Normal prior belief over the set of possible models. This is crucial since it allows us to define **ambiguity shock** as the time-variations in the prior variance over different possible models in a Bayesian fashion. In the real world, whenever there presented increasing doubts, either by investors or by firms, over the underlying models of the aggregate economy, we interpret it as a positive ambiguity shock. Under the smooth model of ambiguity, we highlight the **dual impacts of ambiguity shock** in driving not only confidence but also uncertainty. Widely acknowledged in the decision science, ambiguity aversion results in pessimistic decisions in the sense that agents behave as if they possess a pessimistic belief over the set of possible models. Overweighting pessimistic models when making decisions implies that pessimism over possible models gets amplified in response to a mean-preserving spread in beliefs over possible models, i.e., a positive ambiguity shock. Market confidence drops due to amplified pessimism over models. On the other hand, a positive ambiguity shock makes agents more uncertain over different models. In response, agents rely more on

[3]Agents being ambiguity averse means that they dislike ambiguity more than risk. Put it differently, decision-makers prefer to know the specific probability distribution rather than being ambiguous over it. Ambiguity aversion is natural in the first place since it conceptually differentiates ambiguity from risk and is consistent with many pieces of experimental evidence dating back to the celebrated Ellsberg paradox.

their private information when making expectations manifesting the fact that private information is more valuable in times with heightened "uncertainty". Therefore, belief divergence goes up.

Motivated by this dual impacts of ambiguity shock, we develop a novel theory of ambiguity-driven business cycles that contributes to explain the co-movements across market confidence, belief divergence, and aggregate economy.

Framework and Mechanism. We formalize an otherwise standard real business cycle model that additionally features (a) aggregate demand externalities; (b) ambiguity averse agents (households, workers or firm managers) and finally (c) incomplete information over ambiguous aggregate fundamentals.

In the model, firms differ in productivities, which consists of an aggregate component as well as an idiosyncratic component. The average productivities of all firms are ambiguous in the sense that the cross-sectional mean of idiosyncratic components can be anything on the real-line. Firms have perfect knowledge over own productivity but incomplete information over the ambiguous average productivities of all other firms. The key thing here is that when making output decision, firms need to make expectations over the average productivities of the other firms since it is the sufficient statistics of own demand conditions under aggregate demand externalities. Since firms are operated for the interests of the ambiguity averse households, firms behave as if they are ambiguity averse by themselves. Therefore, output decisions are made as if firms possess a pessimistic belief over models when making expectations over demand conditions. Pessimism at the firm level means that models with a lower demand on average are perceived to be more likely than optimistic models with a higher demand on average. In response to a positive ambiguity shock, firms become more uncertain across different possible models of demand conditions. Such a mean-preserving spread of belief over models amplifies the degree of pessimism of firms resulting in a depressed expectation over own demand conditions. At the aggregate level, there would be an increase in the economy-wide degree of pessimism, which is interpreted as depressed **market confidence**.

Also, firms rely on the observation of own productivity form expectations over demand. This is because individual firm productivity equals to the average productivities of all other firms plus an idiosyncratic shock. In this sense, the observation of own productivity serves as a private information over average productivities of all other firms or own demand conditions. When forming expectation over demand, firms face trade-offs between the use of prior and private information. To note that, we formulate ambiguity shock as the shock to the variance of prior belief over possible models. A positive ambiguity shock results in a less informative prior information over demand conditions from the perspective of firms. Hence it incentivizes more use of private information in forming expectations over demand. Therefore, firm-level output responds more to private information, which further implies that aggregate output responds more to average productivities of all firms. From the perspective of the professional forecasters who are required to submit a personal forecast over aggregate output, this indicates that there are more to be estimated. Also, following the same arguments as in the case of firms, professional forecasters would also use more of their private information to make forecast over the average productivities of the economy. In combine, output forecast of professional forecasters respond more to private information, which heightens **belief divergence**.

Finally, firms hire less labor and cut down output in response to depressed expectation over own demand conditions, which eventually leads to a drop in aggregate output. From this perspectives, ambiguity shock in this paper is nothing more than a particular formulation of **aggregate demand shock**.

To note that, incomplete information is crucial in driving all these results in our model. Technically, this is because if the information is complete, all agents including firms and households can perfectly coordinate not only their beliefs but also their actions. Common knowledge of the economy implies no heterogeneity in beliefs hence zero belief divergence. Also, all firms will have not only perfect knowledge of own productivity but also perfect knowledge of own demand conditions. Therefore, there exists no room for ambiguity shock having any impact on market confidence, belief divergence or aggregate economy. The broader insights are that incomplete information helps us to accommodate a situation where ambiguity is mostly over others' productivity rather than own productivity. Alternatively, ambiguity is mostly over the short-run outlooks of the economy rather than long-run perspectives, which is crucial in understanding the aggregate demand shock nature of ambiguity shock.

Results. The paper starts with a simple business cycle model that abstracts out capital accumulation that allows us to deliver a couple of analytical results that clarifies the main mechanism of the paper. We demonstrate that a positive ambiguity shock generates lower market confidence hence lower aggregate output and larger belief divergence if agents inside the economy are ambiguity averse and the information is incomplete. Therefore, we conclude that fluctuations in market confidence, belief divergence and aggregate economy are the many shades of ambiguity shock.

At the core of these results are the interplay between incomplete information and dual impacts of ambiguity shocks. On the one hand, a positive ambiguity shock makes agents (firms and households) believe that the aggregate fundamental becomes more volatile and on the other hand, aggregate fundamental is turning bad when they are ambiguity averse. With the existence of incomplete information, the former provides increased incentives for the use of private information both in making forecasts and decisions resulting in heightened belief divergence and the latter translates into increased pessimism over aggregate demand resulting in depressed market confidence. Aggregate output plummets since all firms believe their demand is turning bad and in response cut down output.

We deliver such dual impacts of ambiguity shock by a game theoretic interpretation of the equilibrium, which resembles the idea in Angeletos and La'O (2009) such that any business cycle models with incomplete information can be transformed into a beauty contest. This game-theoretic interpretation allows us to clarify the role of incomplete information in driving aggregate fluctuations. It turns out that incomplete information acts as an amplification mechanism for ambiguity. In the extreme case when there presents complete information, ambiguity shock has no impact at the aggregate level. Finally, to note that our aggregate demand channel relies on incomplete information but not on any forms of nominal rigidities. Therefore, aggregate fluctuations generated by ambiguity shock are consistent with the fact that most of the aggregate fluctuations observed in the US data are disconnected from productivity or inflation[4]. Our novel theory of ambiguity-driven business cycles features non-inflationary demand-driven aggregate fluctuations.

We further conduct a couple of quantitative evaluations of the impacts of ambiguity shock within the dynamic RBC framework. Ambiguity shock is shown to be able to generate aggregate co-movements in output, consumption, hours and investment without commensurate movements in labor productiv-

[4]Angeletos, Collard, and Dellas (2016a) identifies a business cycle factor, which is a one-dimensional summary of aggregate movements. The business cycle factor turns out to disconnect from technology or inflation. Further see Gali (1999) and Basu, Fernald, and Kimball (2006) for the former point. And Mavroeidis, Plagborg-Møller, and Stock (2014) for a survey of the empirical evidence on NKPC that corresponds to the latter point, namely inflation puzzle. Also, see Beaudry and Portier (2014) for a couple of evidence that US business cycles are mainly non-inflationary demand-driven.

ity similar to confidence shock alias Angeletos, Collard, and Dellas (2016b) and Huo and Takayama (2015). This further allows us to interpret ambiguity shock as aggregate demand shock quantitatively. What drive the co-movements in quantities are fluctuations in the degree of pessimism over the short run outlooks of the economy, i.e., fluctuations in market confidence due to ambiguity shock. The model is capable of generating empirically plausible countercyclical labor wedge. In this sense, we can alternatively interpret ambiguity shock as the counter-cyclical tax on labor supply or pro-cyclical subsidy on labor demand. Moreover, quantitatively the model can capture cyclical behaviors in cross-sectional dispersions in output forecasts indicated by the SPF dataset. Finally, the estimated market confidence closely tracks Sentiment Index in Michigan Survey of Consumer manifesting the fact that our theory captures movements in market confidence quantitatively.

Contributions. The contributions of our paper are four folds. First of all, we propose a theory of ambiguity-driven business cycles that is capable of generating non-inflationary demand-driven aggregate fluctuations. We complement the current literature with an alternative formulation of aggregate demand shock. Our paper further extends the conventional wisdom in understanding the co-movements across confidence, uncertainty, and aggregate economy by arguing that these can be the endogenous outcomes of ambiguity shock not only qualitatively but also quantitatively.

Secondly, we also contribute to the ambiguity literature with a Bayesian formulation of ambiguity shock based on the smooth model of ambiguity. Conceptually, it differs with the existing literature in the sense that it is a shock to the amount of ambiguity rather than a shock to agents' taste over ambiguity as in Bhandari, Borovicka, and Ho (2016) or a mix of both as in Ilut and Schneider (2014). The unique insight for our Bayesian formulation of ambiguity shock that cannot be shared with the others is that it induces endogenous fluctuations in measured uncertainty, such as dispersion measures. Our paper further contributes to the literature with a particular channel to let second-moment shock, i.e., ambiguity shock, have the first-moment impact. Such a channel relies on the interplay between ambiguity aversion and incomplete information rather than the non-convex adjustment costs as in the theory of uncertainty shock.

Finally, we make one methodological contribution by the provision of a linkage between (a) business cycle models featuring ambiguity aversion and incomplete information and (b) games of incomplete information with ambiguity aversion. The game theoretic interpretation of the business cycle model allows us to build the key economic intuition behind main mechanisms of our paper and to deliver more insights into the interplay between ambiguity, ambiguity aversion, and incomplete information. It turns out incomplete information acts as an amplification mechanism for ambiguity when agents are ambiguity averse.

Layout. The rest of the paper is organized as follows. Section 2 reviews the literature. Section 3 sets up the simple model without capital. Section 4 characterizes the equilibrium by a formal definition as well as a set of optimality conditions. Then in Section 5, impacts of ambiguity shocks are closely studied within the simple model without capital. Section 6 sets up a dynamic RBC model where a couple of quantitative evaluations are conducted. Finally, Section 7 concludes.

2.2. Related Literature

This paper is related to the literature of expectation-driven business cycles including (a) the news shock literature, Beaudry and Portier (2004,2006) and also Jaimovich and Rebelo (2009); (b) noise shock literature, Lorenzoni (2009), Barsky and Sims (2012) and Blanchard, L'Huillier, and Lorenzoni (2013); (c) confidence shock literature, Angeletos and La'O (2013), Angeletos, Collard, and Dellas (2016b) and Huo and Takayama (2015); (d) misspecification shock literature, Bhandari, Borovicka, and Ho (2016); (e) non-bayesian ambiguity shock literature, Ilut and Schneider (2014) and Ilut and Saijo (2016) and finally (f) uncertainty shock literature, Bloom (2009), Bidder and Smith (2012) and Bloom, Floetotto, Jaimovich, Saporta-Eksten, and Terry (2016). We contribute this line of literature with an alternative "exotic" shock, i.e., Bayesian formulation of ambiguity shock, which generates not only aggregate fluctuations but also endogenous co-movements across market confidence and belief divergence within the lens of the business cycles.

Our theory of ambiguity-driven business cycles, on the one hand, directly connects to the theory of confidence shock alias Angeletos and La'O (2013), Angeletos, Collard, and Dellas (2016b) and Huo and Takayama (2015) in generating animal-spirit-like aggregate fluctuations, at the core of which is the interplay between ambiguity aversion and incomplete information. In response to a positive ambiguity shock, firms behave as if their beliefs over average productivities of the others are depressed but beliefs over own productivity, on which they have perfect information, are unaffected. In this sense, we provide a microfoundation of the heterogeneous prior setup in Angeletos, Collard, and Dellas

(2016b). More importantly, our theory of ambiguity shock has the additional ability to capture fluctuations in belief divergence.

On the other hand, our paper also relates to the theory of uncertainty shock alias Bloom (2009) and Bloom, Floetotto, Jaimovich, Saporta-Eksten, and Terry (2016) in generating time-varying measures of uncertainty, specifically for belief divergence. Under our Bayesian formulation, the nature of ambiguity shock is a second-moment "model uncertainty" shock, a positive realization of which incentivizes more use of private information for agents when making output forecasts hence heightened uncertainty. In the broader context, we share the same research spirit with uncertainty shock literature in identifying a particular mechanism that enables second-moment shocks to have first-moment impacts. Uncertainty shock literature relies on non-convex adjustment costs while we rely on the interplay between ambiguity aversion and incomplete information. However, as noted by Angeletos, Collard, and Dellas (2016b), uncertainty shocks as in Bloom (2009) and Bloom, Floetotto, Jaimovich, Saporta-Eksten, and Terry (2016) "can only generate realistic aggregate co-movements through inducing strong pro-cyclical movements in aggregate TFP". From this perspective, it is an alternative formulation of aggregate productivity shock. While in our paper, ambiguity shock is by nature aggregate demand shock. A notable exception in the uncertainty shock literature is Bidder and Smith (2012), who demonstrates that interaction between robust preference and stochastic volatility generates animal spirits fluctuations. We differentiate with them in the underlying preference structure and also with the transmission mechanism from animal spirit to aggregate fluctuations. They rely on news shock channel while we rely on the confidence channel.

There are some other works that study the implications of ambiguity aversion in the context of business cycle models, but with the different mathematical representation of the preferences. For example, Ilut and Schneider (2014) and Bhandari, Borovicka, and Ho (2016). Apart from the above-mentioned difference in the ability to capture fluctuations in belief divergence or measured uncertainty in general, our paper differs with both of the above-mentioned works in a few other aspects. Ilut and Schneider (2014) uses multiple priors preference axiomatized by Gilboa and Schmeidler (1989) to model ambiguity aversion, and ambiguity shock is modeled in a classical statistics fashion. Bhandari, Borovicka, and Ho (2016) uses robust preference proposed by Hansen and Sargent (2001a, 2001b) and focus on time-varying concerns for model misspecification, which can be understood as time-variations in the degree of ambiguity aversion. While in our paper, ambiguity aversion is modeled by (recursive) smooth model of ambiguity axiomatized by Klibanoff, Marinacci, and Mukerji (2005, 2009) and learning over models features smooth rule updating proposed by Hanany and Klibanoff (2009) to ensure dynamic consistency. Our Bayesian formulation of ambiguity shock is a pure shock to the amount of ambiguity rather than a shock to agents' taste over ambiguity as in Bhandari, Borovicka, and Ho (2016) or a mix of both as inIlut and Schneider (2014). Our concept of ambiguity shock is consistent with empirical studies on the recent financial crisis that demonstrate that sudden increases in credit spreads observed during the 2007-2008 crisis are mainly due to the increase in the amount of ambiguity rather than the increase in agents' taste over ambiguity[5].

A notable exception in this area is Ilut and Saijo (2016). In their paper, ambiguity averse preference is represented by the multiple priors preference axiomatized by Epstein and Schneider (2003). Ambiguity aversion is modeled as an amplification mechanism of business cycles rather than devices for exogenous variations. In their paper, recessions are periods of less learning. With an exogenous entropy constrain, reduced learning translates into a larger range of models. Therefore, there features endogenous counter-cyclical ambiguity. They focus on the aggregate implications of their model and succeed in capturing aggregate fluctuation in the data. However, their model implications for dispersion measures are ambiguous. Less learning in recession implies less use of private information when making forecasts. In most of the cases, this implies lower cross-sectional dispersions either in output forecasts or realized outputs. We differ with them in generating the right co-movement pattern in cross-sectional dispersion measures.

Ambiguity averse preferences, especially recursive multiple prior preferences, have been intensively used in the literature to generate asymmetric responses of aggregate variables to aggregate shocks in recessions and boom. For example Epstein and Schneider (2008), Ilut (2012), Ilut, Kehrig, and Schneider (2016), Baqaee (2017) and Zhang (2017). These papers assume there presents ambiguity over precisions of various sources of information to provide agents with state-dependent subjective belief over precisions of related information. Then negative realizations of shocks are associated with a subjective belief of higher precisions hence increased responses to shocks in bad times. In our paper, ambiguity is over the first-moment of shocks but ambiguity shock is of second-moments, which is the unique feature of the smooth model of ambigu-

[5]See Boyarchenko (2012) for more details.

ity. The interplay between ambiguity aversion and incomplete information in our paper also generates counter-cyclical responses to aggregate (productivity) shock. This is because a positive ambiguity shock incentivizes the use of private information in making expectations hence in making output decision. Eventually, aggregate output responds more to aggregate productivity shock. In sum, as far as we know, we are the only paper in the literature that studies joint implications of ambiguity aversion and incomplete information within the lens of business cycles.

Our paper also relates to the theory of beauty contest pioneered by Morris and Shin (2002) then extended by Angeletos and Pavan (2007). We further extend the theory by allowing for ambiguity averse preference. Finally, in a broader context, our paper is also related to those studying coordination games with model uncertainty, such as Chen, Lu, and Suen (2016) and Chen and Suen (2016).

2.3. The Simple Model without Capital

In this section, we construct a static general equilibrium model in the vein of Angeletos and La'O (2009). The model embeds three key features in an otherwise standard real business cycle environment: (a) aggregate demand externalities, (b) incomplete information over the ambiguous aggregate state of the economy and finally (c) smooth model of ambiguity together with ambiguity shock. We first describe the physical environment along with the uncertainty structure and the evolving of information sets of all agents. We then specify the preferences and interim belief systems. Finally, we close up this section by a couple of remarks and interpretations of the setup.

2.3.1. Physical Environment, Shocks and Information Structure

Geography, markets and timing. The economy consists of a continuum of islands, indexed by $j \in J = [0,1]$ and a mainland. On each island j, there exists a continuum of firms, indexed by $(i,j) \in I \times J = [0,1]^2$ and a continuum of workers, indexed by $(m,j) \in M \times J = [0,1]^2$. Island firms and workers interact with each other in the locally competitive labor market for the production of differentiated island commodities indexed by j. These commodities are traded in a centralized market operated on the mainland, where a continuum of consumers, indexed by $h \in H = [0,1]$ and a large number of competitive final good producers inhabit. We assume that consumer h and a continuum of workers $\{(h,j); j \in J\}$ constitute a large household indexed by $h \in H$, who owns a continuum of firms $\{(h,j);j \in J\}$. By doing so, we ensure the existence of a representative household at the mainland and a continuum of representative firms and workers on every island. This is because, as it will become evident later, there exist no heterogeneities, either in fundamental or in information, within an island.

We focus on the static setup in this section. There is only one period, say period t, which is decomposed into three stages. At stage zero, period t shocks are realized. At stage 1, island-specific competitive labor markets open up. The representative household sends out workers to each islands. On island j, firms make labor demand decisions and symmetrically, workers make labor supply decisions on the basis of incomplete information over the ambiguous aggregate state of the economy. At stage 2, on the mainland, the centralized commodities market opens up. All uncertainty, either risk or ambiguity, is resolved. Final good producers produce. And the representative household makes consumption decisions upon receiving all the transfers from workers and firms on the basis of perfect information.

Households. The utility of the representative household is given by:

$$\frac{C_t^{1-\gamma}-1}{1-\gamma}-\chi\int_J \frac{N_{j,t}^{1+\epsilon}}{1+\epsilon}dj$$

where γ is the relative risk aversion and ϵ is the inverse Frisch elasticity of labor supply. To note that, in our static setup, γ also controls for the income effects of labor supply. The corresponding budget constraint of household is such that

$$P_tC_t=\int_J W_{j,t}N_{j,t}dj+\int_J\int_I \Pi_{i,j,t}d\left(i,j\right)$$

where P_t denotes the price of final goods, $\int_J W_{j,t}N_{j,t}dj$ denotes the total labor income and finally $\int_J\int_I \Pi_{i,j,t}d\left(i,j\right)$ denotes the total realized firm profits.

Island firms. Island j firms use labor only for the production of island j commodity. The production function of firm (i,j) is given by

$$Y_{j,t}=A_{j,t}N_{i,j,t}^{1-\alpha} \tag{2.1}$$

where $A_{j,t}$ is the island-specific productivity and the realized profit is given by

$$\Pi_{i,j,t}=P_{j,t}Y_{i,j,t}-W_{j,t}N_{i,j,t}$$

where $W_{j,t}$ denotes the nominal wage on island j in period t and $P_{j,t}$ denotes the market price of island j commodity to be determined at stage 2 when the centralized markets opens up. Since it is assumed that it is the representative household who owns the firm, any realized profits are to be transferred to the consumer for the purchase of final goods for consumption. Therefore, in the absence of any uncertainty concerns, island j firms care about the consumer valuation over its profits given by

$$\frac{u'(C_t)}{P_t}\Pi_{\iota,j,t}$$

where P_t is the price of final goods normalized to 1.

Final-good producers. The competitive final-good sector employs a CES production technology given by

$$Y_t = \left(\int_J Y_{j,t}^{\frac{\theta-1}{\theta}} dj\right)^{\frac{\theta}{\theta-1}}$$

where θ is elasticities of substitution among island commodities that controls the strength of aggregate demand externalties. Demand function for island j commodity is, therefore, given by

$$Y_{j,t} = \left(\frac{P_{j,t}}{P_t}\right)^{-\theta} Y_t$$

where $P_t \equiv \left(\int_J P_{j,t}^{1-\theta} di\right)^{\frac{1}{1-\theta}}$ denotes the price of final goods normalized to 1.

Productivity and ambiguity shocks. Aggregate productivity $a_t \equiv \log A_t$ follows a Normal distribution with mean 0 and variance σ_ζ^2

$$a_t \sim N\left(0, \sigma_\zeta^2\right)$$

Island-specific productivity, defined by $a_{j,t} \equiv \log A_{j,t}$, equals to aggregate productivity plus some idiosyncratic productivity shock $\iota_{j,t}$:

$$a_{j,t} = a_t + \iota_{j,t}$$

Idiosyncratic productivity shocks $\iota_{j,t}$ are assumed to be i.i.d normally distributed with mean ω_t and variance σ_ι^2. Objectively, the cross-sectional mean of idiosyncratic productivity shock is zero, i.e. $\omega_t = 0$. However, agents inside the economy cannot fully "understand" it. They possess some ambiguity over it. Specifically, they believe that anything on the real-line can be a potential candidate for ω_t. And they possess a common zero-mean[6] Normal prior belief over $\omega_t \in \mathcal{R}$:

$$\omega_t \sim N\left(0, e^{\psi_t}\right) \quad \text{with}$$

where ψ_t measures the amount of ambiguity perceived by the agents[7]. We assume that ψ_t is chosen by nature such that

$$\psi_t = \overline{\psi} + \tau_t \quad \text{with} \quad \tau_t \sim N\left(0, \sigma_\tau^2\right) \tag{2.2}$$

where $\overline{\psi}$ denotes the amount of ambiguity perceived by all agents at the ambiguous steady state[8]. And we interpret τ_t as ambiguity shock, which is Normally distributed with mean 0 and variance σ_τ^2.

Information structure. Denote $\mathcal{I}_{t,0}$, $\mathcal{I}_{j,t,1}$ and $\mathcal{I}_{t,2}$ as the information sets that are available to all agents at stage 0 of period t, are only available to island j agents at stage 1 of period t and are available to all agents at stage 2 of period t, respectively. We define these information sets by

$$\mathcal{I}_{t,0} = \{\psi_t\} \qquad \mathcal{I}_{j,t,1} = \mathcal{I}_{t,0} \cup \left\{a_{j,t}\right\} \qquad \mathcal{I}_{t,2} = \cup_j \mathcal{I}_{j,t,1} \cup \{\zeta_t\} \tag{2.3}$$

There are a couple of implicit assumptions behind this information structure. First of all, information is symmetric within each island but is asymmetric across islands. Secondly, ambiguity shock τ_t happens at the beginning of each period t. Thirdly, at stage 1 of period t, island j productivity $a_{j,t}$ is only accessible for island j agents. In this sense, $a_{j,t}$ serves as the private information of island j agents over the average productivity $\int_J a_{j,t} dj$, which is ambiguous. Therefore, labor supply and demand decisions on each island are made under incomplete information over the ambiguous aggregate state of the economy. Fourthly, $\mathcal{I}_{t,2}$ contains the complete set of local information that is originally dispersed at stage 1. This is justifiable since commodities prices or transfers perfectly reveal the island-specific productivity that is previously dispersed at stage 1. Fifthly,

[6]The objective model ω_t are assumed to be inside the set of possible models of all agents. By assuming this, we rule out any mis-specification concerns and focus on ambiguity. See Peter Hansen and Marinacci (2016) for a detailed discussion of the differences between mis-specification and ambiguity

[7]Maccheroni, Marinacci, and Ruffino (2013) proposes to use the variance of the ex-ante expected utility of a particular model to quantify the amount of ambiguity in general information structure, which is shown to be consistent with a quadratic approximation akin to Arrow-Pratt approximation. Our measure of the amount of ambiguity is consistent with theirs ordinally under Normality.

[8]Ambiguous steady state refers to the state the economy converges to in the absence of any shocks but taking into account of the existence of ambiguity.

all uncertainty, either risk (state uncertainty over a_t) or ambiguity (model uncertainty over ω_t) is resolved at stage 2 of period t. Hence consumption decisions are made under perfect information.[9]

We close the description of the physical environment, shocks and information structure by the timeline of our model in Figure 2.1.

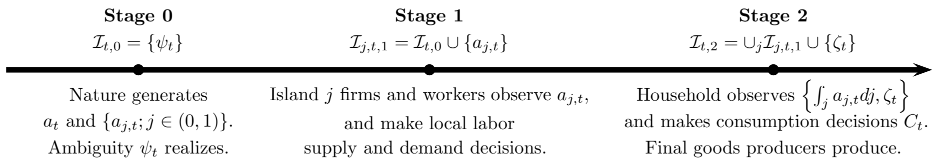


Figure 2.1. *Timeline for Period t*

2.3.2. Preferences and Interim Belief Systems

All agents inside the economy are ambiguity averse in the sense that they dislike ambiguity more than risk. Ambiguity averse preferences can be mathematically represented by the smooth model of ambiguity first proposed by Klibanoff, Marinacci, and Mukerji (2005), which features a separation between the amount of ambiguity (characteristics of subjective belief) and degree of ambiguity aversion (characteristics of decision makers' tastes). Such a separation is appealing because only with it, we can seriously study the impacts of ambiguity shock upon controlling for decision makers' taste over it. It is widely acknowledged that Bayesian updating is only dynamic consistent under expected utility preferences. To ensure dynamic consistency *across stages* within a period, we employ the smooth rule of updating, which is a re-weighted Bayesian updating rule proposed in Hanany and Klibanoff (2009)[10].

At stage 2, when all uncertainty is resolved, the representative household's preferences are represented by a standard utility function. In what follows, we carefully specify the preferences and interim belief systems for the representative household at stage 1 of period t and formulate the stage 1 workers' and firms' problems when there features incomplete information is over the ambiguous fundamentals.

[9]The assumption that all period t uncertainty resolves at the second stage of period t is in some sense ad-hoc to ensure tractability. However, most of the key messages delivered in this paper do not rely on this particular assumption on information structure.

[10]See Appendix for a discussion over dynamic consistency

Preference of the representative household at stage 1. At stage 1, preference of the representative household is represented by the smooth model of ambiguity proposed in Klibanoff, Marinacci, and Mukerji (2005). The corresponding island j workers' problem is such that

$$\max_{N_{j,t}} \int_{\mathcal{R}} \phi \left(E^{\omega_t}_{j,t,1} \left[\frac{C_t^{1-\gamma} - 1}{1-\gamma} - \chi \int_J \frac{N_{j,t}^{1+\epsilon}}{1+\epsilon} dj \right] \right) \widetilde{f}^h_{j,t,1} (\omega_t) \, d\omega_t \tag{2.4}$$

$$\text{s.t.} \quad P_t C_t = \int_J W_{j,t} N_{j,t} dj + \int_J \int_I \Pi_{i,j,t} d(i,j) \tag{2.5}$$

where the island j workers decide how much labor $N_{j,t}$ to supply into island-specific competitive labor market given nominal wage $W_{j,t}$.

Here $\phi(x)$ is some strictly increasing and concave function, whose curvature captures decision makers' taste over ambiguity, i.e., degree of ambiguity aversion and $E^{\omega_t}_{j,t,1}[\cdot]$ denotes the mathematical expectation conditioned on $\mathcal{I}_{t,1}$ under a particular model ω_t for the cross-sectional mean of idiosyncratic productivity shock ω_t. And $\widetilde{f}^h_{j,t,1}(\omega_t)$ stands for the probability density function of interim belief of island j workers over period t ambiguity ω_t. It is nothing more than the posterior belief over possible models ω_t, which follows smooth rule of updating

$$\widetilde{f}^h_{j,t,1}(\omega_t) \propto \underbrace{\frac{\phi'\left(E^{\omega_t}_{j,t,0} \left[\frac{C_t^{1-\gamma}-1}{1-\gamma} - \chi \int_J \frac{N_{j,t}^{1+\epsilon}}{1+\epsilon} dj \right] \right)}{\phi'\left(E^{\omega_t}_{j,t,1} \left[\frac{C_t^{1-\gamma}-1}{1-\gamma} - \chi \int_J \frac{N_{j,t}^{1+\epsilon}}{1+\epsilon} dj \right] \right)}}_{\text{Weights}} \underbrace{f\left(a_{j,t} | \omega_t\right) f_t(\omega_t)}_{\text{Bayesian Kernel}} \tag{2.6}$$

Here $f\left(a_{j,t}|\omega_t\right)$ is the conditional probability density function of $a_{j,t}$ under a particular model ω_t, which is the Normal density with mean ω_t and variance $\sigma_\zeta^2 + \sigma_\iota^2$, and $f_t(\omega_t)$ stands for the period t prior belief density over ω_t, which is the Normal density with mean 0 and variance e^{ψ_t}.

Relative to the standard Bayesian updating, the smooth rule puts more weight to the model that provides higher marginal incentive to act ex-ante (at stage 0) when comparing to its ex-post (at stage 2) counterparts:

$$\phi'\left(E^{\omega_t}_{t,0} \left[\frac{C_t^{1-\gamma} - 1}{1-\gamma} - \chi \int_J \frac{N_{j,t}^{1+\epsilon}}{1+\epsilon} dj \right] \right) > \phi'\left(E^{\omega_t}_{t,2} \left[\frac{C_t^{1-\gamma} - 1}{1-\gamma} - \chi \int_J \frac{N_{j,t}^{1+\epsilon}}{1+\epsilon} dj \right] \right)$$

Through re-weighting in interim belief, incentives to act in ex-ante (at stage 0) and in ex-post (at stage 2) can be aligned with each other, which leads to dynamic consistency across stages within a period.

Firm problem at stage 1. Firm problem can be formulated as follows:

$$\max_{N_{i,j,t}} \int_{\mathcal{R}} \phi \left(E^{\omega_t}_{j,t,1} \left[\frac{C_t^{-\gamma}}{P_t} \left(P_{j,t} Y_{i,j,t} - W_{j,t} N_{i,j,t} \right) \right] \right) \widetilde{f}^{f}_{j,t,1} \left(\omega_t \right) d\omega_t \tag{2.7}$$

subject to the production function of island j firms (2.1). The firm decides how much labor to hire by taking nominal wage $W_{j,t}$ as given and by making expectation over its terms of trade $P_{j,t}$ to be determined at stage 2 by

$$Y_{j,t} \equiv \int_I Y_{i,j,t} di = \left(\frac{P_{j,t}}{P_t} \right)^{-\theta} Y_t \tag{2.8}$$

Interim belief systems of island j firms follows an extended smooth rule of updating given by

$$\widetilde{f}^{f}_{j,t,1} \left(\omega_t \right) \propto \underbrace{\frac{\phi' \left(E^{\omega_t}_{j,t,0} \left[\frac{C_t^{1-\gamma}-1}{1-\gamma} - \chi \int_J \frac{N_{j,t}^{1+\epsilon}}{1+\epsilon} dj \right] \right)}{\phi' \left(E^{\omega_t}_{j,t,1} \left[\frac{C_t^{-\gamma}}{P_t} \left(P_{j,t} Y_{j,t} - W_{j,t} N_{j,t} \right) \right] \right)}}_{\text{Weights}} \underbrace{f \left(a_{j,t} | \omega_t \right) f_t \left(\omega_t \right)}_{\text{Bayesian Kernel}} \tag{2.9}$$

Unlike the standard smooth rule of updating, firms' incentives to act are not aligned in a pure ex-ante versus ex-post sense. Instead, the proposed extended smooth rule of updating aligns incentives to act of the representative household in ex-ante with that of the firms in ex-post. This ensures *dynamic consistency from the perspective of the representative household*. That is if we allowed the household to make ex-ante contingency plans of production for island firms, the contingency plans are to be respected ex-post by firms when it is their turn to move.

To note that, we formulate the firms' problem in a way that the firms are ambiguity averse by themselves. We can justify the formulation of firms' problem by arguing that firms' are maximizing the shareholder value. Therefore, firms behave as if they are ambiguity averse by themselves and share the same belief with their shareholders when evaluating the marginal benefit of labor demand. The additional concavity introduced by the ϕ function manifests the former point, and the extended smooth rule of updating takes care the latter.

Therefore, we can alternatively formulate the firms' problem by

$$\max_{N_{i,j,t}} \int_{\mathcal{R}} E_{j,t}^{\omega_t} \left[SDF_t \left(P_{j,t} Y_{i,j,t} - W_{j,t} N_{i,j,t} \right) \right] f_{j,t,1} \left(\omega_t \right) d\omega_t \tag{2.10}$$

where the stochastic discount factor SDF_t is given by

$$SDF_t \equiv \phi' \left(E_{j,t,0}^{\omega_t} \left[\frac{C_t^{1-\gamma} - 1}{1-\gamma} - \chi \int_J \frac{N_{j,t}^{1+\epsilon}}{1+\epsilon} dj \right] \right) \frac{C^{-\gamma}}{P_t}$$

Here the stochastic discount factor takes cares not only households' risk attitude $\frac{C^{-\gamma}}{P_t}$ but also ambiguity attitude $\phi' \left(E_{j,t,0}^{\omega_t} \left[\frac{C_t^{1-\gamma}-1}{1-\gamma} - \chi \int_J \frac{N_{j,t}^{1+\epsilon}}{1+\epsilon} dj \right] \right)$. The two formulations (2.7) and (2.10) are isomorphic to each other.

To close up the description of the model, we assume that $\phi(x)$ takes the constant absolute ambiguity aversion (CAAA) form for simplicity and tractability:

Assumption 2.1. (CAAA) We assume $\phi(x) = -\frac{1}{\lambda} e^{-\lambda x}$ where $\lambda \geq 0$ measures degree of ambiguity aversion of all agents.

2.3.3. Remarks and Interpretations

We conclude this section by some remarks and interpretations on the three key features of our model that have been listed at the beginning of this section.

1. Ambiguity is introduced into the model in the form of the cross-sectional mean of idiosyncratic productivity shock. But this does not mean there is any ambiguity over local economic conditions. Instead, firms and workers on island j have the perfect understanding of own island productivity, but an incomplete and ambiguous understanding of average productivity of all other islands $\int_J a_{j,t} dj$. This is because from the perspective of island j agents, cross-sectional mean of idiosyncratic productivity shock ω_t can be regarded as some temporary aggregate productivity shocks. Therefore, if local economic decisions are made solely depending on expectations over local economic conditions, output or labor will not respond to ambiguity shock at all[11]. This is the exact reason why we need aggregate demand externalities.

2. In our model, with aggregate demand externalities, incomplete and ambiguous information over the average productivity of all other islands $\int_J a_{j,t} dj$

[11] There are still some inter-temporal impact of ambiguity shock on consumption-saving trade-offs. But this would, in general, imply no aggregate co-movements, which is the sentence to death for any business cycle model.

can be translated into incomplete and ambiguous information over own demand conditions. We show in later sections that under the smooth model of ambiguity and the smooth rule of updating, fluctuations in the amount of ambiguity, i.e., ambiguity shock τ_t, generate fluctuations in island j agents' belief over local demand conditions, which eventually maps into aggregate fluctuations. In this sense, we can formally interpret our ambiguity shock as a particular formulation of aggregate demand shocks. This differentiates our paper with Ilut and Schneider (2014), where ambiguity shocks are designed to be some form of news shocks about future productivity.

3. Ambiguity shock τ_t in our model is, by design, a second-moment shock. This is the primary reason we can generate fluctuations in belief divergence, i.e., cross-sectional dispersion of ex-ante output forecast. The key question here is whether or not such a second-moment shock can generate the first-moment impact at the aggregate level. The answer is yes if we have the smooth model of ambiguity together with the smooth rule of updating. From this perspective, we share the same spirit with Bloom (2009) and Bloom, Floetotto, Jaimovich, Saporta-Eksten, and Terry (2016) in identifying possible mechanisms that enable the second-moment shocks to have first-moment impacts[12].

2.4. Equilibrium Characterization

In this section, we first define the equilibrium of the model and then derive a set of optimality conditions that jointly describe the equilibrium allocations and beliefs of all agents. Finally, we demonstrate how to characterize the *conditional log-normal equilibrium* associated with these optimality conditions.

2.4.1. Equilibrium Definition

Definition 2.1 (Equilibrium). *An equilibrium consists of a set of*

- *allocations* $\left\{ \{N_{i,j,t}, Y_{i,j,t}\}_{(i,j)\in J}, \{N_{j,t}\}_{j\in J}, Y_t \right\}$;
- *factors and commodities prices* $\left\{ \{W_{j,t}\}_{j\in J}, \{P_{j,t}\}_{j\in J}, P_t \right\}$;
- *information sets* $\left\{ \mathcal{I}_{t,0}, \{\mathcal{I}_{j,t,1}\}_{j\in J}, \mathcal{I}_{t,2} \right\}$;
- *exogenous shocks* $\left\{ \tau_t, \zeta_t, \{\iota_{j,t}\}_{j\in J} \right\}$;

[12]In Bloom (2009) and Bloom, Floetotto, Jaimovich, Saporta-Eksten, and Terry (2016), the existence of non-convex adjustment costs generates the real value of wait-and-see when "uncertainty" (risk in its nature) goes up, which has first moment impact on firms hiring and investment decisions.

- *and finally interim beliefs over the set of possible models* $\left\{\widetilde{f}^{h}_{j,t,1}(\omega_t), \widetilde{f}^{f}_{j,t,1}(\omega_t)\right\}_{j\in J}$

such that:

- *Information sets* $\left\{\mathcal{I}_{t,0}, \{\mathcal{I}_{j,t,1}\}_{j\in J}, \mathcal{I}_{t,2}\right\}$ *are defined in (2.30)*
- *At stage* 1*, given factors and commodities prices* $\left\{\{W_{j,t}\}_{j\in J}, \{P_{j,t}\}_{j\in J}, P_t\right\}$ *and the interim beliefs over the set of possible models* $\left\{\widetilde{f}^{h}_{j,t,1}(\omega_t), \widetilde{f}^{f}_{j,t,1}(\omega_t)\right\}_{j\in J}$, $\{N_{j,t}\}$ *solves the workers' problem (2.4) and* $\{N_{i,j,t}, Y_{i,j,t}\}_{(i,j)\in J}$ *solves the firms' problem (2.7)*
- *Interim beliefs are such that:* $\widetilde{f}^{h}_{j,t,1}(\omega_t)$ *is given by (2.6) and* $\widetilde{f}^{f}_{i,j,t,1}(\omega_t)$ *is given by (2.9).*
- *Market clears for island-specific labor markets*

$$\int_I N_{i,j,t} di = N_{j,t}$$

- *Market clears for island commodities*

$$\int_I Y_{i,j,t} dj \equiv Y_{j,t} = \left(\frac{P_{j,t}}{P_t}\right)^{-\theta} Y_t \tag{2.11}$$

with

$$Y_t = \left(\int_J Y_{j,t}^{\frac{\theta-1}{\theta}} dj\right)^{\frac{\theta}{\theta-1}}$$

where the price of final goods $P_t \equiv \left(\int_J P_{j,t}^{1-\theta} di\right)^{1/(1-\theta)}$ *is normalized to* 1.

- *Market clears for final good*

$$Y_t = C_t$$

2.4.2. Optimality Conditions

We can characterize the equilibrium with a set of optimality conditions. Detailed derivations can be found in Appendix.

First of all, within island j labor market, optimal labor supply is governed by the following condition

$$\chi N_{j,t}^{\epsilon} = W_{j,t} \int_{\mathcal{R}} E_{j,t,1}^{\omega_t} \left[u' \left(C_t \right) \right] \widetilde{f}_{j,t,1} \left(\omega_t \right) d\omega_t \tag{2.12}$$

Workers on island j equate stage 1 valuation of the marginal benefit of labor (RHS) with marginal disutility of labor. On the other side of the labor market, optimal labor demand condition is given by

$$W_{j,t} \int_{\mathcal{R}} E_{j,t,1}^{\omega_t} \left[u' \left(C_t \right) \right] \widetilde{f}_{j,t,1} \left(\omega_t \right) d\omega_t = \left(\int_{\mathcal{R}} E_{j,t,1}^{\omega_t} \left[u' \left(C_t \right) P_{j,t} \right] \widetilde{f}_{j,t,1} \left(\omega_t \right) d\omega_t \right) \left((1 - \alpha) \frac{Y_{j,t}}{N_{j,t}} \right) \tag{2.13}$$

Firms on island j equate stage 1 valuation of the marginal cost of labor (LHS) with the marginal benefit (RHS). Unlike standard expected utility preferences, ambiguity aversion implies that when evaluating marginal effects at stage 1 of period t, firms and workers on island j employs a distorted posterior belief over the set of possible models given by

$$\widetilde{f}_{j,t,1} \left(\omega_t \right) \propto \underbrace{\phi' \left(E_{j,t,0}^{\omega_t} \left[\frac{C_t^{1-\gamma} - 1}{1 - \gamma} - \chi \int_J \frac{N_{j,t}^{1+\epsilon}}{1 + \epsilon} dj \right] \right)}_{\text{Belief Distortion}} \underbrace{f \left(a_{j,t} | \omega_t \right) f_t \left(\omega_t \right)}_{\text{Bayesian Kernel}} \tag{2.14}$$

It tells that whenever a model ω_t generates lower ex-ante (stage 0) expected utility for the representative household at the margin of ω_t, island j firms and workers tend to regard it as the more likely one in their posteriors. Put it differently, ambiguity aversion implies a pessimistic posterior belief over the set of possible models. To note that, firms on island j have the same distorted posterior belief over the set of possible models as island j workers, which is the by-product of the extended smooth rule we assumed in the interim belief systems of island firms.

Combing (2.12) and (2.13), equilibrium allocation of labor can be summarized by the following key equation for labor market:

$$\chi N_{j,t}^{\epsilon} = \left(\underbrace{\int_{\mathcal{R}} E_{j,t,1}^{\omega_t} \left[u' \left(C_t \right) \left(\frac{Y_{j,t}}{Y_t} \right)^{-\frac{1}{\theta}} \right] \widetilde{f}_{j,t,1} \left(\omega_t \right) d\omega_t}_{\text{marginal utility of island } j \text{ commodity}} \right) \left(\underbrace{(1 - \alpha) \frac{Y_{j,t}}{N_{j,t}}}_{\text{marginal productivity}} \right) \tag{2.15}$$

The LHS of this key equation is the marginal disutility of labor, and the RHS is the multiplication of (a) marginal utility of island j commodity and (b) marginal productivity of labor. The equation simply says, in the labor market equilibrium, stage 1 valuation of private benefit of labor equates with the private cost of labor. Similar condition also appears in Angeletos and La'O (2009) and Angeletos, Iovino, and La'O (2016). There are two main differences between ours and theirs. First of all, there is another round of integration over models due to the existence of ambiguity. Second, agents use a distorted posterior belief due to ambiguity aversion.

2.4.3. Joint Approximation of Allocation and Belief

Using island production function $Y_{j,t} = A_{j,t} N_{j,t}^{1-\alpha}$ and the market clearing condition for final goods $Y_t = C_t$, we can transform (2.15) into a fixed point condition over allocation $\left\{Y_{j,t}\right\}_{j \in J}$:

$$\chi Y_{j,t}^{\frac{1+\epsilon}{1-\alpha}-1+\frac{1}{\theta}} = (1-\alpha) A_{j,t}^{\frac{1+\epsilon}{1-\alpha}} \left(\int_{\mathcal{R}} E_{j,t,1}^{\omega_t} \left[Y_t^{\frac{1}{\theta}-\gamma} \right] \widetilde{f}_{j,t,1}(\omega_t) \, d\omega_t \right) \tag{2.16}$$

where the distorted posterior belief $\widetilde{f}_{j,t,1}(\omega_t)$ is given by

$$\widetilde{f}_{j,t,1}(\omega_t) \propto \underbrace{\phi' \left(E_{j,t,0}^{\omega_t} \left[\frac{Y_t^{1-\gamma}-1}{1-\gamma} - \chi \int_J \frac{\left(Y_{j,t}/A_{j,t}\right)^{(1+\epsilon)/(1-\alpha)}}{1+\epsilon} dj \right] \right)}_{\text{Belief Distortion}} \underbrace{f\left(a_{j,t}|\omega_t\right) f_t(\omega_t)}_{\text{Bayesian Kernel}} \tag{2.17}$$

To ensure complementarity in productions across islands, we make the following parametric restriction for the simple model without capital[13]:

Assumption 2.2. (Complementarity) It is assumed that $\frac{1}{\theta} > \gamma$ when there is no capital.

Increase in output of all other islands $k \neq j \in J$, on the one hand, raises the demand for island j commodities due to aggregate demand externalities, but on the other hand generates upward pressure on the wage rate of island j due to income effect of labor supply. Assumption 2.2 ensures that income effect of labor

[13] To see why this is the case, observe that under perfect information, (2.16) can be simplified into

$$\chi Y_{j,t}^{\frac{1+\epsilon}{1-\alpha}-1+\frac{1}{\theta}} = \left(\frac{\eta-1}{\eta}\right)(1-\alpha) A_{j,t}^{\frac{1+\epsilon}{1-\alpha}} Y_t^{\frac{1}{\theta}-\gamma}$$

It is straight-forward to show $\partial Y_{j,t}/\partial Y_t > 0$ if and only if $\frac{1}{\theta} - \gamma > 0$.

supply is so weak that the channel of aggregate demand externalities dominates that of the income effects. Later in Section 2.6, we drop this assumption when there exists capital accumulation and assume $\theta = 1$, i.e., a Cobb-Douglas technology for the aggregation of island output $\left\{Y_{j,t}\right\}_{j\in J}$ and $\gamma = 1$, i.e., log utility for consumption. Such a parameterization also ensures complementarity in production given the existence of consumption smoothing incentive, which weakens income effect of labor supply.

The technical complication here is that the distorted posterior belief $\widetilde{f}_{j,t,1}\left(\omega_t\right)$ is not orthogonal to equilibrium allocation. Allocations and beliefs have to be solved simultaneously in equilibrium. As a result, the equilibrium of the economy is the solution to a double fixed point conditions: one solves (2.16) characterizing the equilibrium cross-sectional allocation $\left\{Y_{j,t}\right\}_{j\in J}$ conditional on any stage 1 distorted posterior belief over possible models $\widetilde{f}_{j,t,1}\left(\omega_t\right)$ and the other one solves (2.17) characterizing equilibrium stage 1 distorted posterior belief over the set of possible models conditional on any cross cross-sectional allocation $\left\{Y_{j,t}\right\}_{j\in J}$ of the economy.

Definition 2.2 (Conditional Log-Normal Equilibrium). *An allocation $\left\{Y_{j,t}, Y_t\right\}_{j\in J}$ constitutes a conditional Log-Normal equilibrium if both $Y_{j,t}|\psi_t$ and $Y_t|\psi_t$ are Log-Normally distributed.*

Lemma 2.1. Up to second order, distorted posterior belief over the set of possible models $\widetilde{f}_{j,t,1}\left(\omega_t\right)$ can be Normally approximated if allocation $\left\{Y_{j,t}, Y_t\right\}_{j\in J}$ constitutes a conditional Log-Normal equilibrium.

Proof. See Appendix. ∎

Lemma 2.2. Allocation $\left\{Y_{j,t}\left(a_{j,t}, \psi_t\right)\right\}_{j\in J}$ constitutes a conditional Log-Normal equilibrium if distorted posterior belief over possible models $\widetilde{f}_{j,t,1}\left(\omega_t\right)$ is Normal.

Proof. Directly follows Angeletos and La'O (2009). ∎

The complication of the joint determination (or approximation) of allocations and beliefs can be greatly simplified once we narrow our analysis down to the focus of conditional log normal equilibrium defined in Definition 2.2. On the one hand, conditional Log-Normal equilibrium embeds the standard

Log-Normal equilibrium or log-linearized equilibrium as a special case when there is no ambiguity shock. While on the other hand, it can be justified, up to an approximation sense, by Lemma 2.1 and Lemma 2.2 in a self-fulfilling fashion. The following proposition characterizes the approximated conditional log-normal equilibrium.

Proposition 2.1 (Equilibrium Characterization)**.** *Under some regularity conditions, there exists a unique* ***approximated*** *symmetric conditional Log-Normal equilibrium where the allocation* $\left\{Y_{j,t}, Y_t\right\}_{j\in J}$ *is such that*

$$y_{j,t} \equiv \ln Y_{j,t} = \left(\underbrace{y^* + \overline{h}_y\left(\overline{\psi}\right)}_{\text{Ambiguous SS}}\right) + \underbrace{\kappa_{ya_j}\left(\psi_t, \lambda\right) \cdot a_{j,t}}_{\text{Use of Private Info.}} + \underbrace{\widehat{h}_y\left(\psi_t, \lambda\right)}_{\text{Impact of Amb. Shock}} \tag{2.18}$$

and

$$y_t \equiv \ln Y_t = \left(\underbrace{y^* + \overline{h}_y\left(\overline{\psi}\right)}_{\text{Ambiguous SS}}\right) + \underbrace{\kappa_{ya_j}\left(\psi_t, \lambda\right) \cdot \int_J a_{j,t} dj}_{\text{Use of Private Info.}} + \underbrace{\widehat{h}_y\left(\psi_t, \lambda\right)}_{\text{Impact of Amb. Shock}} \tag{2.19}$$

where $y^* + \overline{h}_y\left(\overline{\psi}\right)$ *denotes the ambiguous steady state output. And* $\kappa_{ya_j}\left(\psi_t, \lambda\right)$*, the slope of output w.r.t. productivity, is called the use of private information, which is a function of the amount of ambiguity* ψ_t *and degree of ambiguity aversion* λ*. Finally,* $\widehat{h}_y\left(\psi_t; \lambda\right)$ *denotes the impact of ambiguity shock on output satisfying*

$$\widehat{h}_y\left(\overline{\psi}, \lambda\right) = 0$$

Finally, the distorted posterior belief over the set of possible models is Normal with mean μ_t *and variance* σ_t^2 *such that*

$$\mu_t = \left(\frac{e^{\psi_t} + g_\sigma\left(\psi_t, \lambda\right)}{\sigma_\zeta^2 + \sigma_\iota^2 + e^{\psi_t} + g_\sigma\left(\psi_t, \lambda\right)}\right) x_{j,t} + \left(\frac{\sigma_\zeta^2 + \sigma_\iota^2}{\sigma_\zeta^2 + \sigma_\iota^2 + e^{\psi_t} + g_\sigma\left(\psi_t, \lambda\right)}\right) g_\mu\left(\psi_t, \lambda\right)$$

and

$$\sigma_t^2 = \left(\frac{\sigma_\zeta^2 + \sigma_\iota^2}{\sigma_\zeta^2 + \sigma_\iota^2 + e^{\psi_t} + g_\sigma\left(\psi_t, \lambda\right)}\right)\left(e^{\psi_t} + g_\sigma\left(\psi_t, \lambda\right)\right)$$

where the distortion in mean $g_\mu(\psi_t,\lambda)$ and in variance $g_\sigma(\psi_t,\lambda)$ are given by

$$g_\mu(\psi_t,\lambda) = -\lambda\kappa_{J\omega}\left(\frac{1}{1+\lambda\kappa_{J\omega\omega}e^{\psi_t}}\right)e^{\psi_t} \quad g_\sigma(\psi_t,\lambda) = -\left(\frac{\lambda\kappa_{J\omega\omega}e^{\psi_t}}{1+\lambda\kappa_{J\omega\omega}e^{\psi_t}}\right)e^{\psi_t} \tag{2.20}$$

with $\kappa_{J\omega}$ and $\kappa_{J\omega\omega}$ being functions of κ_{ya_j} such that

$$\kappa_{J\omega} = e^{(1-\gamma)y^*}\left(\left(1+(1-\gamma)\overline{h}_y\right)\kappa_{ya_j} - \left(1+\left(\frac{1+\epsilon}{1-\alpha}\right)\overline{h}_y\right)\left(\kappa_{ya_j}-1\right)\right) > 0 \tag{2.21}$$

$$\kappa_{J\omega\omega} = e^{(1-\gamma)y^*}\left[\left(1+(1-\gamma)\overline{h}_y\right)(1-\gamma)\kappa^2_{ya_j,t} - \left(1+\left(\frac{1+\epsilon}{1-\alpha}\right)\overline{h}_y\right)\left(\frac{1+\epsilon}{1-\alpha}\right)\left(\kappa_{ya_j}-1\right)^2\right] > 0 \tag{2.22}$$

Proof. See Appendix. ∎

The approximation turns out to be quite accurate. In Appendix 2.B, we conduct a numerical check and demonstrate that our approximation is not only accurate but also conservative in the sense it under-estimates the impact of ambiguity shock.

The allocations (2.18) and (2.19) are akin to those of the Angeletos and La'O (2009). Log deviations from the ambiguous SS for island output $\widehat{y}_{j,t} \equiv y_{j,t} - \left(y^* + \overline{h}_y(\overline{\psi})\right)$ and aggregate output $\widehat{y}_t \equiv y_j - \left(y^* + \overline{h}_y(\overline{\psi})\right)$ can be expressed into a linear function of island $a_{j,t}$ and average productivity $\int_j a_{j,t}dj$, respectively. We name the slope of these linear functions $\kappa_{ya_j}(\psi_t,\lambda)$ as the use of private information, which is a function of the amount of ambiguity ψ_t. Furthermore, the intercept term $\widehat{h}_y(\psi_t;\lambda)$ controls the impact of ambiguity shock on aggregate output, given the fact that it is zero when evaluated at the ambiguous steady state. These two terms (a) the use of private information $\kappa_{ya_j}(\psi_t,\lambda)$ and (b) the impact of ambiguity shock on aggregate output $\widehat{h}_y(\psi_t;\lambda)$ are at the core of our analysis. Later in this Section 2.5, we study the impacts of ambiguity shocks through comparative static analysis over these terms and demonstrate how ambiguity shock can possibly generate co-movements across market confidence, belief divergence and aggregate economy.

2.5. Impacts of Ambiguity Shock

In this section, we analyze the impacts of ambiguity shock by conducting a couple of comparative static analysis. We start by providing a game theoretic interpretation of the equilibrium allocations of our business cycle model. Such a game theoretic interpretation clarifies the main mechanisms of our paper, which are the dual impacts of ambiguity shock. We then demonstrate how such dual impacts of ambiguity shock are to be mapped into fluctuations in aggregate output, market confidence, and belief divergence. Finally, we close up this section with some discussions about the interplay between incomplete information and ambiguity aversion.

2.5.1. Game Theoretic Interpretation

To build up the economic intuition behind impacts of ambiguity shock, we present a game theoretic interpretation of the equilibrium of our business cycle model, which resembles the beauty contest in Morris and Shin (2002) and Angeletos and Pavan (2007), but with a distorted information structure that captures the belief distortion or pessimism in belief due to ambiguity aversion.

Proposition 2.2. *The approximated equilibrium allocations* $\left\{Y_{j,t}, Y_t\right\}_{j\in J}$ *are identical to that of a beauty contest such that*

$$y_{j,t} = \kappa_a a_{j,t} + \kappa_y E_{j,t}\left[y_t\right]$$

where the coefficients κ_a *and* κ_y *are such that*

$$\kappa_a = \frac{\frac{1+\epsilon}{1-\alpha}}{\frac{1+\epsilon}{1-\alpha} - 1 + \frac{1}{\theta}} \qquad\qquad \kappa_y = \frac{\frac{1}{\theta} - \gamma}{\frac{1+\epsilon}{1-\alpha} - 1 + \frac{1}{\theta}} \in (0,1)$$

The information structure is distorted such that

$$\begin{aligned}
&\widetilde{a}_{j,t} = \widetilde{a}_t + \widetilde{\iota}_{j,t}, \quad \widetilde{\iota}_{j,t} \sim N\left(0, \sigma_\iota^2\right) \\
&\widetilde{a}_t \sim N\left(g_\mu\left(\psi_t, \lambda\right), \sigma_\zeta^2 + e^{\psi_t} + g_\sigma\left(\psi_t, \lambda\right)\right)
\end{aligned}$$

where distortions $\left\{g_\mu\left(\psi_t, \lambda\right), g_\sigma\left(\psi_t, \lambda\right)\right\}$ *are given by (2.20) and satisfies the followings*

$$g_\mu\left(\psi_t, \lambda\right) \le 0, \quad g_\mu\left(-\infty, \lambda\right) = 0, \quad g_\mu\left(\psi_t, 0\right) = 0, \quad \frac{\partial g_\mu\left(\psi_t, \lambda\right)}{\partial \psi_t} < 0.$$

and

$$g_\sigma\left(\psi_t, \lambda\right) \le 0, \quad g_\sigma\left(-\infty, \lambda\right) = 0, \quad g_\sigma\left(\psi_t, 0\right) = 0, \quad \frac{\partial\left(e^{\psi_t} + g_\sigma\left(\psi_t, \lambda\right)\right)}{\partial \psi_t} > 0$$

Proof. See Appendix. ∎

Within the beauty contest interpretation, island output $y_{j,t}$ is the linear combination of island productivity $a_{j,t}$ and island j expectation of aggregate output because the former controls the marginal cost of production on island j and the latter manifests island j's forecast over own demand conditions. However, unlike standard beauty contest, perceived distribution of aggregate productivity is distorted both in mean and variance due to the existence of ambiguity and ambiguity aversion. Here κ_y corresponds to the notion of coordination motive in the beauty contest literature. Its magnitude $\kappa_y \in (0,1)$ is the product of Assumption 2.2, which ensures complementarity in action and uniqueness in allocations once we fix a distorted information structure. Such a game theoretic interpretation provides us a natural laboratory to study the impacts of ambiguity shock. In what follows, we utilize this beauty contest interpretation and discuss in more details how ambiguity generates fluctuations in market confidence, belief divergence and aggregate economy once and for all.

2.5.2. Market Confidence, Belief Divergence and Aggregate Fluctuations

What are the impacts of ambiguity shocks? As evident from Proposition 2.1, ambiguity shock affects the allocations of the economy by fluctuating decision makers' distorted posterior belief over the set of possible models. Proposition 2.2 further decomposes the impacts of a positive ambiguity shock into two parts: one generates an increased pessimism over aggregate fundamental and the other one creates higher perceived volatility of aggregate fundamentals. We call them the dual impacts of ambiguity shock.

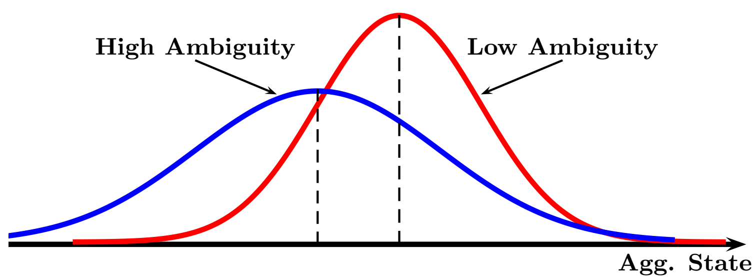


Figure 2.1. *Impact of Ambiguity Shock: Main Mechanism*

Figure 2.1 plots the perceived "as if" distribution of aggregate fundamental for the low level of ambiguity, i.e., ψ_t is small, and for the high level of ambiguity, i.e., ψ_t is large. It is named "as if" because these are the subjective beliefs over aggregate fundamentals that would deliver exactly the same allocation as our baseline model when agents have expected utility preferences. Since island j agents have perfect understanding over own productivity, distorted prior belief over aggregate fundamental $\widetilde{a}_t$ can be translated into distorted prior belief over aggregate demand of island j. Therefore, a positive ambiguity shock makes decision-makers believe that the aggregate demand becomes worse on average and more volatile in their "as if" subjective prior. The former maps into lower output, either island or aggregate, at the margin. While the latter maps into an increased incentive in the use private information when making expectation over aggregate demand hence when making factors demand and supply decisions. We summarize these results in the following proposition

Proposition 2.3. *A positive ambiguity shock that increases the amount of ambiguity ψ_t generates lower aggregate output in the sense that*

$$\frac{\partial \widehat{h}_y\left(\psi_t, \lambda\right)}{\partial \psi_t} < 0 \tag{2.23}$$

if agents are ambiguity averse, i.e. $\lambda > 0$. Moreover, equilibrium use of private information $\kappa_{ya_j}\left(\psi_t, \lambda\right)$ is an increasing function of amount of ambiguity ψ_t:

$$\frac{\partial \kappa_{ya_j}\left(\psi_t, \lambda\right)}{\partial \psi_t} > 0. \tag{2.24}$$

Proof. See Appendix. ∎

At the core of understanding (2.23) is the increased degree of pessimism over the set of possible models. In fact, there are two forces at work that deepen agents' degree of pessimism over the set of possible models, one fundamental and one strategic. A positive ambiguity shock, on the one hand, increases the amount of ambiguity faced by all agents. In response, agents behave more pessimistic. This is the fundamental or direct channel. On the other hand, a positive ambiguity shock induces all other agents to use more of their private information when making output decisions. Under aggregate demand externalities, this raises the amount of ambiguity in firms' demand structure, which further increases the degree of pessimism. This is the strategic or indirect channel. Through the fundamental and strategic channels, a positive ambiguity shock raises all agents' degree of pessimism over the outlooks of the economy, which eventually drives down economic activities.

Market confidence. We define market confidence as the "economy-wide average of agents' first-order expectations of aggregate output of the economy". Such a definition is consistent with the practice of most of the survey exercise. For example, in the Michigan Survey of Consumer, consumers are asked about whether they believe output will go up or down in the following year. In a broader context, it is a short-cut for agents forecasts about the outlooks of the economy.

Definition 2.3 (Market Confidence). *Market confidence is mathematically defined to be*

$$Conf.\left(\psi_t, \lambda\right) \equiv \int_J \int_{\mathcal{R}} E_{j,t}^{\omega_t}\left[y_t\right] \widetilde{f}_{j,t,1}\left(\omega_t\right) d\omega_t dj$$

To note that, market confidence is defined under the distorted belief over the set of possible models $\widetilde{f}_{j,t,1}\left(\omega_t\right)$. By doing so, we implicitly assume that ambiguity averse agents would use their as if belief, i.e., the pessimistic belief to make forecasts. Also, we integrate individual agents' first-order belief over j to be consistent with the idea that market confidence is an economy-wide concern. Following Proposition 2.1, it is directly to have that

$$Conf.\left(\psi_t, \lambda\right) = y^* + \overline{h}_y\left(\overline{\psi}\right) + \kappa_{ya_j}\left(\psi_t, \lambda\right)\left(\frac{\sigma_\iota^2}{\sigma_\zeta^2 + e^{\psi_t} + g_\sigma\left(\psi_t, \lambda\right)}\right) g_\mu\left(\psi_t, \lambda\right) + \widehat{h}_y\left(\psi_t, \lambda\right)$$

Observe that economy-wide average of agents first-order belief over average productivity is given by

$$\int_J \int_{\mathcal{R}} E_{j,t}^{\omega_t}\left[\int_J a_{j,t} dj\right] \widetilde{f}_{j,t,1}\left(\omega_t\right) d\omega_t d = \kappa_{ya_j}\left(\psi_t, \lambda\right)\left(\frac{\sigma_\iota^2}{\sigma_\zeta^2 + e^{\psi_t} + g_\sigma\left(\psi_t, \lambda\right)}\right) g_\mu\left(\psi_t, \lambda\right) \tag{2.25}$$

Being hit by a positive ambiguity shock, all agents become more pessimistic over the aggregate fundamental, i.e., $g_\mu\left(\psi_t, \lambda\right)$ decreases. On the other hand, all agents perceive the aggregate fundamental become more volatile $e^{\psi_t} + g_\sigma\left(\psi_t, \lambda\right)$ increases. The former increases the economy-wide pessimism. But the latter reduces economy-wide pessimism because agents find it optimal to use more of the private information, the economy-wide average of which is objectively zero. However, in equilibrium, the former always dominates the latter implying that all agents are becoming more pessimistic over average productivity. The point is all agents understand that all the others are more pessimistic. And they also understand that the other agents understand this increase in economy-wide pessimism. Further, they all understand that all the others understand that the others understand this, etc. The consequence of such an higher-order thinking over each others eventually leads to a drop in aggregate output, i.e., $\widehat{h}_y(\psi_t, \lambda)$. This further depresses market confidence. Finally, all agents also understand that the others all perceive aggregate fundamental being more volatile. Hence they understand that all the others will use more of their private information. Therefore, they know that aggregate output will respond more to aggregate fundamental, i.e., $\kappa_{ya_j}(\psi_t, \lambda)$ increases. This raises output forecasts' reliance on the pessimistic belief over average productivity, which further depresses market confidence. We summarize this result in the Corollary 2.1.

Corollary 2.1. (Depressed Market Confidence) A positive ambiguity shock that increases the amount of ambiguity perceived by agents depresses market confidence in the sense that

$$\frac{\partial Conf.(\psi_t, \lambda)}{\partial \psi_t} < 0$$

Proof. Directly following Proposition 2.2 and 2.3. ∎

To note that in the model, island j agents have perfect information over own productivity. An increase in the amount of ambiguity depresses island j agents' belief over the average productivity of the other islands without changing beliefs over own productivity. These movements in belief are isomorphic to those of a negative confidence shock under heterogenous prior setup alias Angeletos, Collard, and Dellas (2016b). In this sense, our paper provides an alternative micro-foundation for the heterogeneous prior setup by generating endogenous movements in confidence.

Belief Divergence. To discuss the impact of ambiguity shock on belief divergence, we need to define the expectation formation process of professional forecasters formally. Equivalently, we ask the following question: what are professional forecasters' attitudes towards ambiguity?

Figure 2.2 plots realized real GDP growth rates and its average year-ahead forecasts from Survey of Professional Forecasters. There displays no significant pessimism over time in SPF forecasts. To be consistent with this observation, we assume that professional forecasters are ambiguity neutral. If not, ambiguity aversion would imply a systematic pessimism in average forecasts due to

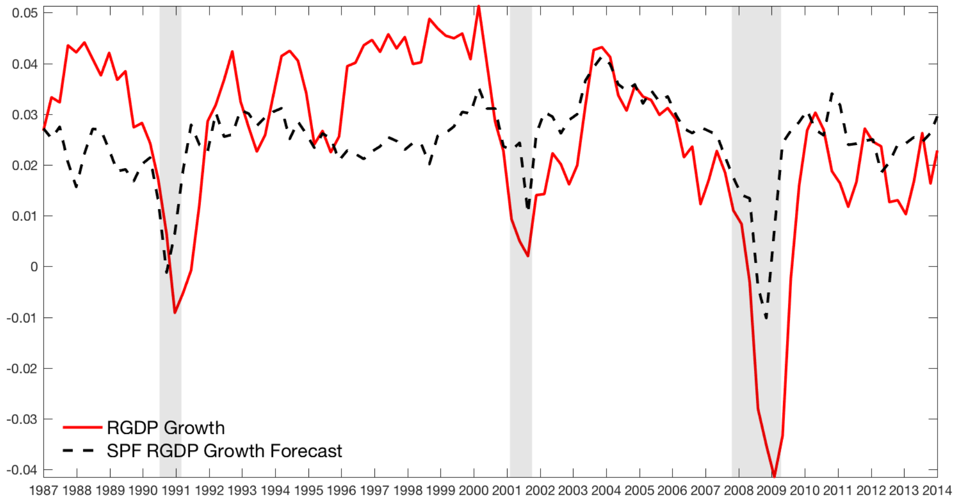


Figure 2.2. *SPF Forecasts of Real GDP Growth Rate.*
Note: The figure plots real GDP growth rate (red line) and year-ahead real GDP growth rate forecast by professional forecasters (dashed black line) between 1987Q1 and 2014Q4. Forecast data is from Survey of Professional Forecasters provided by the Federal Reserve Bank of Philadelphia and real GDP growth data is from Saint-Louis Federal Reserve Economic Database.

distorted belief over the set of possible models.

Assumption 2.3. Professional forecasters possess ambiguity over the cross-sectional mean of idiosyncratic productivity shock. However, they are ambiguity neutral, i.e., $\lambda = 0$.

Furthermore, each island j is assumed to inhabit a professional forecaster indexed by $j \in J = [0,1]$, who submits his forecast about period t aggregate output y_t at stage 1 of period t. In this section, we assume that professional forecaster j shares the same information as island j agents. Hence his forecast over aggregate output can be expressed into

$$E_{j,t}\left[y_t\right] = y^* + \overline{h}_y\left(\overline{\psi}\right) + \underbrace{\kappa_{ya_j}\left(\psi_t,\lambda\right)}_{\uparrow}\underbrace{\left(\frac{\sigma_\zeta^2 + e^{\psi_t}}{\sigma_\zeta^2 + e^{\psi_t} + \sigma_\iota^2}\right)}_{\uparrow} a_{j,t} + \widehat{h}_y\left(\psi_t,\lambda\right). \tag{2.26}$$

Therefore, belief divergence measured by the cross-sectional dispersion of ex-ante output forecast is given by

$$FD_t(\psi_t,\lambda) \equiv \int_J \left(E_{j,t}[y_t] - \overline{E}_t[y_t]\right)^2 dj = \kappa_{ya_j}^2(\psi_t,\lambda)\left(\frac{\sigma_\zeta^2 + e^{\psi_t}}{\sigma_\zeta^2 + e^{\psi_t} + \sigma_\iota^2}\right)^2 \sigma_\iota^2 \quad (2.27)$$

Corollary 2.2 summarizes the impacts of ambiguity shocks on belief divergence.

Corollary 2.2. (Heightened Belief Divergence) Therefore, a positive ambiguity shock that increases the amount of ambiguity ψ_t raises belief divergence in the sense that

$$\frac{\partial FD_t(\psi_t,\lambda)}{\partial \psi_t} > 0. \quad (2.28)$$

Proof. Straight-forward following the intuition below. ■

The intuition here is straight-forward. A positive ambiguity shock makes firms and workers on all islands believe, in their "***as if***" subjective prior, that the aggregate fundamental is more volatile, which increases the incentive to use private information when forming expectations over own demand conditions. This maps into increased responsiveness of island output $y_{j,t}$ to island productivity $a_{j,t}$ because it is $a_{j,t}$ that serves as the private information over aggregate demand for island j agents. This raises cross-sectional dispersion of island output. Upon aggregation, we also have aggregate output y_t responds more to average productivity $\int_J a_{j,t}dj$.

From the perspective of professional forecaster j, increase in κ_{ya_j} implies that "there are more to estimate". Moreover, when he estimates the average productivity $\int_J x_{j,t}dj$, he tends to rely more on his private information $a_{j,t}$ in the sense that $\frac{\partial\left(\frac{\sigma_\zeta^2+e^{\psi_t}}{\sigma_\zeta^2+e^{\psi_t}+\sigma_\iota^2}\right)}{\partial\psi_t} > 0$. This is because he believes, in his "as if" subjective prior, the aggregate fundamental is now more volatile. These two in combine increase the responsiveness of forecaster j's forecast to private information $a_{j,t}$, which eventually leads to higher cross-sectional dispersion in output forecasts ex-ante, i.e., an increase in belief divergence. To note that, the economy itself does not become more dispersed or more volatile. It is the increased responsiveness to idiosyncratic shocks that drives up the cross-sectional dispersion. This differentiates our paper with the theory of uncertainty shock as in Bloom

(2009) and Bloom, Floetotto, Jaimovich, Saporta-Eksten, and Terry (2016), who take fluctuations in dispersion as model inputs rather than model output.

We provide a summary of impacts of ambiguity shock by combing Proposition 2.3, Corollary 2.1 and Corollary 2.2:

Proposition 2.4 (Impacts of Ambiguity Shock). *If decision makers are ambiguity averse, i.e., $\lambda > 0$, a positive ambiguity shock that increases the amount of ambiguity ψ_t generates*

- *lower market confidence;*
- *larger belief divergence;*
- *and finally lower aggregate output on average.*

Proof. Directly following Proposition 2.3, Corollary 2.1 and Corollary 2.2. ∎

2.5.3. Discussion: incomplete information and ambiguity aversion

In 2.5.2, we highlight the strategic force behind the impacts of ambiguity shocks in driving aggregate fluctuations. This is closely related to incomplete information embedded in our paper. We close up the analysis of Section 2.5 by some additional discussions about the interplay between incomplete information and ambiguity based on the game theoretic interpretation developed in Section 2.5.1. By varying degree of incompleteness in information, impacts of uncertainty shocks are studied.

Incomplete information is of primary importance. When information is complete, i.e., $\sigma_\iota^2 = 0$, island j agents face no uncertainty regarding decisions of agents on other islands. Then all agents will have the perfect understanding of the whole economy. Therefore, ambiguity shocks play no role in driving aggregate fluctuations without incompleteness in information. Corollary 2.3 summarizes this result.

Corollary 2.3. Ambiguity shocks have no impacts on aggregate output $\widehat{h}_y(\psi_t, \lambda) = 0$ when information is complete $\sigma_\iota^2/\sigma_\zeta^2 = 0$.

Proof. Straight-forward following the proof of Proposition 2.1. ∎

To note that, completeness in information only requires that there exists no private information. It does not necessarily imply perfect information[14]. Corollary 2.3 still holds when there presents information friction as long it is common across agents. Unlike our simplified case with $\sigma_\iota^2/\sigma_\zeta^2 = 0$, when this is the case, there still presents the distorted posterior belief over the set of possible models. However, all agents have common knowledge over Y_t under complete information. Then (2.16) transforms into

$$\chi Y_{j,t}^{\frac{1+\epsilon}{1-\alpha}-1+\frac{1}{\theta}} = (1-\alpha)\, A_{j,t}^{\frac{1+\epsilon}{1-\alpha}} Y_t^{\frac{1}{\theta}-\gamma}$$

which leaves no room for ambiguity or ambiguity shock to have any impacts. To interpret, it is the imperfect coordination across islands, possibly due to imperfect communication as in Angeletos and La'O (2013), that matters rather than imperfect information.

The question of interest naturally arises. Does more incompleteness in information amplify or dampen the impacts of ambiguity shocks? We answer these questions by varying not only σ_ι^2 but also σ_ζ^2 while holding the amount of ambiguity constant at the A-SS level.

There are two forces at work. On the one hand, more incompleteness in information, i.e., larger σ_ι^2 or lower σ_ζ^2, reduces the incentive in the use of private information because private information is becoming less informative not only about the aggregate state but also about the models. This marginally changes degree of pessimism since $g_\mu\left(\overline{\psi},\lambda\right)$ is a function of both $\kappa_{J\omega}\left(\overline{\kappa}_{ya_j},\overline{h}_y\right)$ and $\kappa_{J\omega\omega}\left(\overline{\kappa}_{ya_j},\overline{h}_y\right)$, which are functions of $\overline{\kappa}_{ya_j}$. It can be proved that mean distortion $g_\mu\left(\overline{\psi},\lambda\right)$ marginally decreases w.r.t degree of information incompleteness.

Lemma 2.3. Mean distortion at the A-SS $g_\mu\left(\overline{\psi},\lambda\right)$ marginally decreases in σ_ι^2 and increases in σ_ζ^2:

$$\frac{\partial g_\mu\left(\overline{\psi},\lambda\right)}{\partial \sigma_\iota^2} < 0 \qquad\qquad \frac{\partial g_\mu\left(\overline{\psi},\lambda\right)}{\partial \sigma_\zeta^2} > 0$$

Proof. See Appendix. ∎

[14]See Angeletos and Lian (2016) for more formal discussions.

On the other hand, fixing the belief distortion $g_\mu(\psi_t, \lambda) = g_\mu^*$ and $g_\mu(\psi_t, \lambda) = g_\sigma^*$, more incompleteness in information implies more load of belief distortion on island and aggregate output. To see this, express island output $y_{j,t}$ into an infinite sum over a complete hierarchy of all higher-order beliefs of aggregate fundamentals $\widetilde{a}_t$:

$$y_{j,t} = \kappa_a \left(a_{j,t} + \sum_{n=1}^{\infty} \kappa_y^n E_{j,t}^n [\widetilde{a}_t] \right) \tag{2.29}$$

where higher-order beliefs of island j agents are defined recursively such that

$$E_{j,t}^1 [\widetilde{a}_t] \equiv E_{j,t} [\widetilde{a}_t] \qquad E_{j,t}^n [\widetilde{a}_t] \equiv E_{j,t} \left[\int_J E_{j,t}^{n-1} [\widetilde{a}_t] \, dj \right] \quad \forall n \geq 1$$

It turns out that increased incompleteness in information increases the response of any order of belief to prior information. Observe that it is the prior of aggregate fundamental being distorted. Therefore, it must be the case that all orders of belief are exposed to more belief distortion. Therefore, incompleteness in information naturally amplify the impact of ambiguity shock at the margin when we fix the belief distortion. The above two forces interact with each other, which indicates that incomplete information is actually an amplifying mechanism for the impact of ambiguity. Proposition 2.5 summarizes this result.

Proposition 2.5. *At the ambiguous steady state (A-SS), a marginal increase in σ_ι^2 that increases the degree of information incompleteness, decreases the use of private information and amplifies the impact of ambiguity:*

$$\frac{d\overline{\kappa}_{ya_j}}{d\sigma_\iota^2} < 0 \qquad \frac{d\overline{h}_y}{d\sigma_\iota^2} < 0$$

On the contrary, a marginal increase in σ_ι^2 that reduces the degree of information incompleteness, increases the use of private information and dampens the impact of ambiguity:

$$\frac{d\overline{\kappa}_{ya_j}}{d\sigma_\zeta^2} > 0 \qquad \frac{d\overline{h}_y}{d\sigma_\zeta^2} > 0$$

Proof. See Appendix. ∎

2.5.4. A summary

Taking stock, this section has shown that a positive ambiguity shock generates a recession with depressed market confidence and heightened belief divergence qualitatively. In what follows, we study the impacts of ambiguity shocks quantitatively within an extended dynamic RBC model. We demonstrate that ambiguity shock in our theory can generate reasonable co-movements in quantities at the aggregate level. The availability of our theory is quantitatively evaluated by bringing the observable implications into the data.

2.6. The Dynamic RBC Model: Quantitative Evaluation

In this section, we illustrate the quantitative potential of our theory by studying a dynamic RBC model. We first set up the model. We then move to the discussion about our quantitative methodology in the estimation of conditional Log-Normal equilibrium (Definition 2.2). Finally, observable implications of ambiguity shock, for both the aggregate quantities, belief divergence and market confidence, are then assessed through a calibrated version of the model.

2.6.1. Model Setup

Geography, markets and timing. The economy consists of a continuum of islands, indexed by $j \in J = [0,1]$ and a mainland. On each island j, there exist a continuum of firms, indexed by $(i,j) \in I \times J = [0,1]^2$ and a continuum of workers, indexed by $(m,j) \in M \times J = [0,1]^2$. Firms on island j hire labor and capital from locally competitive factor markets for the production of island-specific commodity j. These commodities are traded in a centralized market operated on the mainland, where a continuum of consumers, indexed by $h \in H = [0,1]$ and a large number of final good producers inhabit. We assume that consumer h and a continuum of workers $\{(h,j); j \in J\}$ constitute a large household indexed by $h \in H$, who owns a continuum of firms $\{(h,j); j \in J\}$. As in the case of the simple model, our model admits a representative household at the mainland and a continuum of representative firms and workers on every island.

Time is discrete, indexed by $t \in \{0,1,...\}$ and each period t is decomposed into three stages. At stage zero, period t shocks are realized. At stage 1, island competitive factor markets open up. Island j firms make labor and capital demand decisions and symmetrically, the representative household sends out workers to each island, who make labor supply decisions on the basis of incomplete information over the ambiguous concurrent aggregate state of the economy. At stage 2, on the mainland, the centralized commodities market opens up. All uncertainty, either risk or ambiguity, over the concurrent aggregate state of the economy resolved. Final goods producers produce. And the representative household makes consumption and saving decisions upon receiving, capital income, labor income and all the transfers from island firms upon perceiving ambiguity over the future aggregate state of the economy. Here we assume that it is the representative household who owns the capital, therefore, saving takes the form of island-specific investments for all islands. Therefore, the capital supply of island j in period $t+1$ is pre-determined at stage 2 of period t by the representative household.

Households. Period utility of the representative household is given by:

$$u\left(C_t\right) - \chi \int_J \frac{N_{j,t}^{1+\epsilon}}{1+\epsilon} dj$$

where ϵ is the inverse Fisher elasticity of labor supply. Therefore, the flow budget constraint is such that

$$P_t C_t + P_t \int_J K_{j,t+1} - (1-\delta) K_{j,t} dj = \int_J W_{j,t} N_{j,t} dj + \int_J R_{j,t} K_{j,t} dj + \int_J \Pi_{j,t} dj$$

where $\int_j R_{j,t} K_{j,t} dj$ and $\int_J W_{j,t} N_{j,t} dj$ denote the total capital and labor income of all islands respectively and $\int_J \Pi_{j,t} dj$ is the transfers of realized profits from all island firms.

Island firms. Island j firms use labor and capital for the production of island j commodity. The production function is Cobb-Douglas:

$$Y_{j,t} = A_{j,t} N_{j,t}^{1-\alpha} K_{j,t}^{\alpha}$$

where $A_{j,t}$ is the island-specific productivity and the realized profit is given by

$$\Pi_{j,t} = P_{j,t} Y_{j,t} - W_{j,t} N_{j,t} - R_{j,t} K_{j,t}$$

where $W_{j,t}$ and $R_{j,t}$ denote the competitive factors prices on island j in period t and $P_{j,t}$ denotes the market price of island j commodity in period $t$to, which is determined at stage 2 when the centralized markets for commodities open up. Since it is the large representative household who owns the firm, any realized profits are to be transferred to the household for the purchase of final goods for consumption and investment. Therefore, in the absence of any uncertainty concerns, island j firms care about the consumer valuation over their profits given by

$$\frac{u'(C_t)}{P_t}\Pi_{j,t}$$

where P_t is the aggregate price level to be normalized to 1.

Productivity and ambiguity shocks. Aggregate productivity $a_t \equiv \log A_t$ follows an AR(1) process

$$a_t = \rho a_{t-1} + \zeta_t$$

where ζ_t is the aggregate productivity shock in period t that follows a Normal distribution with mean 0 and variance σ_ζ^2.

Island-specific productivity, defined by $a_{j,t} \equiv \log A_{j,t}$, equals to aggregate productivity plus an idiosyncratic productivity shock $\iota_{j,t}$:

$$a_{j,t} = a_t + \iota_{j,t}$$

Idiosyncratic productivity shocks $\iota_{j,t}$ are assumed to be i.i.d normally distributed with mean ω_t and variance σ_ι^2. Objectively, the cross-sectional mean of idiosyncratic productivity shocks are zero for all periods, i.e. $\omega_t = 0\ \forall t > 0$. However, agents inside the economy cannot fully understand it. Instead, they possess some ambiguity over the complete set of cross-sectional means of idiosyncratic productivity shocks, i.e., $M \equiv \{\omega_t : \forall t \geq 0\}$.

At the very beginning of time, say period 0, all agents subjectively believe that all $\omega_t \in M$ are i.i.d Normally distributed with mean 0 and variance $\overline{\sigma}_\omega^2 \equiv e^{\overline{\psi}}$. Here $\overline{\sigma}_\omega^2$ or $\overline{\psi}$ measures the amount of ambiguity agents possess in the A-SS. Ambiguity in the past does not last forever. As it will become evident later, concurrent ambiguity is resolved at stage 2 of that period. Therefore, at stage 0 of any period t, agents inside the economy only possess ambiguity over concurrent and future cross-sectional means of idiosyncratic productivity shocks, i.e., $M_t \equiv \{\omega_{t+k} : \forall k \geq 0\}$. The amount of ambiguity they possess at this point, denoted by ψ_t, is time-varying and governed by an AR(1) process

$$\psi_t = \left(1 - \rho_\psi\right)\overline{\psi} + \rho_\psi \psi_{t-1} + \tau_t$$

where τ_t is the ambiguity shock assumed to be Normally distributed with mean 0 and variance σ_τ^2. We close the description of the ambiguity process by explicitly specifying the common subjective prior belief of all agents over $M_t =$

$\{\omega_{t+k} : \forall k \geq 0\}$ at stage 0 of period t:

$$\omega_{t+k} \sim i.i.d\ N\left(0, e^{\psi_{t,t+k}}\right) \quad \forall k \geq 0$$

where the amount of ambiguity agents perceived over ω_{t+k} at period t, denoted as $\psi_{t,t+k}$, is an increasing affine function of ψ_t:

$$\psi_{t,t+k} = \left(1 - \rho_\psi^k\right)\overline{\psi} + \rho_\psi^k \psi_t$$

This particular structure for prior beliefs ensures prior consistency over the future ambiguity $\omega_{t+k}, \forall k \geq 1$. It simply tells that period t prior over future ambiguity ω_{t+k} meets period $t+k$ prior over ω_{t+k} if there are no more ambiguity shocks in between. Put it differently,

$$\phi_{t,t+k} = E_t\left[\psi_{t+k}\right]$$

Ambiguity shock τ_t in the process of ambiguity ψ_t, by its nature, can be understood a changing prior process. Also, we implicitly assume that a positive ambiguity shock in period t, i.e., $\tau_t > 0$, makes agents become "more ambiguous" [15] over all the future ambiguity ω_{t+k}. However, it is period t biased in the sense that it raises concurrent ambiguity more than future ambiguity and the increase in ambiguity is mean-reverting such that for ambiguity ω_{t+k} in the very far future $k \to +\infty$, the subjective belief stays in its A-SS belief, i.e. $\lim_{k\to+\infty} \psi_{t,t+k} = \overline{\psi}$.

Information structure. Denote $\mathcal{I}_{t,0}$, $\mathcal{I}_{j,t,1}$ and $\mathcal{I}_{t,2}$ as the information sets that are available to all agents at stage 0 of period t, are only available to island j agents at stage 1 of period t and are available to all agents at stage 2 of period t, respectively. Recursively, we can define these information sets by

$$\mathcal{I}_{t,0} = \mathcal{I}_{t-1,2} \cup \{\psi_t\} \qquad \mathcal{I}_{j,t,1} = \mathcal{I}_{t,0} \cup \left\{a_{j,t}\right\} \qquad \mathcal{I}_{t,2} = \cup_j \mathcal{I}_{j,t,1} \cup \{\zeta_t\} \tag{2.30}$$

To note that all concurrent uncertainty, either risk (state uncertainty over a_t) or ambiguity (model uncertainty over ω_t) are resolved at stage 2 of period t. Hence consumption-saving decisions by households are made upon perceiving ambiguity over the future outlooks of the economy only. Also because idiosyncratic productivity shocks are i.i.d, $\left\{a_{j,t}\right\}_{j\in J}$ tells no more information than $\int_j a_{j,t} dj$ does regarding island j productivity in period $t+1$. Therefore, we can simplify information set at stage 2 of period t by $\mathcal{I}_{t,2} = \mathcal{I}_{t,0} \cup \left\{\int_j a_{j,t} dj, \omega_t\right\}$. To simplify

[15] Here more ambiguous means an increase in the amount of ambiguity perceived by all agents.

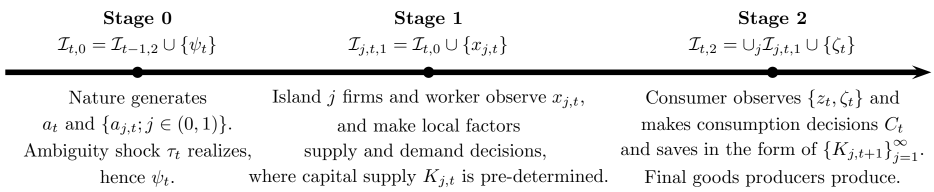


Figure 2.1. *Timeline for Period t*

notation, we further transform the information structure into

$$\mathcal{I}_{t,0} = \mathcal{I}_{t-1,2} \cup \{\psi_t\} \qquad \mathcal{I}_{j,t,1} = \mathcal{I}_{t,0} \cup \{x_{j,t}\} \qquad \mathcal{I}_{t,2} = \mathcal{I}_{t,0} \cup \{z_t, \zeta_t\}$$

where $x_{j,t} \equiv \zeta_t + \iota_{j,t}$ denotes the de-facto private information at stage 1 over aggregate productivity shocks and $z_t = \zeta_t + \omega_t$ denotes the de-facto public information at stage 2 over aggregate productivity shocks. Figure 2.1 displays the timeline and information sets for period t in our dynamic RBC model.

Preference of the representative household at stage 2. Denote $s_{t+1} \equiv \mathcal{I}_{t+1,2} \backslash \mathcal{I}_{t,2}$ as the arrival of new information at stage 2 between two consecutive periods t and $t+1$. We summarize belief of the representative household at stage 2 by two corresponding Bayesian posteriors: (a) $\pi_M(s_{t+1}|\mathcal{I}_{t,2})$, the Bayesian posterior of s_{t+1} at stage 2 of period t under a particular model M, and (2) $\mu(M|\mathcal{I}_{t,2})$, the Bayesian posterior over the entire the set of possible models $M \in \mathcal{M}$.

Preference of the representative household at stage 2 of period t, therefore, can be represented by the recursive smooth model of ambiguity proposed by Klibanoff, Marinacci, and Mukerji (2009):

$$V_t(C; \mathcal{I}_{t,2}) = u(C_t) + \underbrace{\beta\phi^{-1}\left(\int_{\mathcal{M}} \phi\left(\int_{\mathcal{S}_{t+1}} V_{t+1}(C; \mathcal{I}_{t,2}, s_{t+1})\, d\pi_M(s_{t+1}|\mathcal{I}_{t,2})\right) d\mu(M|\mathcal{I}_{t,2})\right)}_{\text{Utility Equivalent of the Ambiguous Continuation Value}}$$

where $\phi(x)$ is some strictly increasing and concave function, whose curvature captures decision makers' taste for ambiguity, i.e., the degree of ambiguity aversion.[16]

Can learning over time resolve all ambiguity in the long run? Klibanoff, Marinacci, and Mukerji (2009) proves that if ϕ^{-1} is Lipschitz and the space of

[16]Theorem 3 in Klibanoff, Marinacci, and Mukerji (2009) proves that under some regularity conditions, there are unique and monotonic V_t. For most of business cycle applications, those regularity conditions can be easily satisfied.

ambiguous model parameters is finite, the recursive smooth model of ambiguity will converge uniformly to expected utility preferences with true model parameters. In our model, agents are ambiguous over an infinite parameter space $\omega \equiv \{\omega_t : \forall t \geq 0\}$. This prevents ambiguity to vanish in the long run through learning.

Moreover, the additional concavity in $\phi(x)$ is to capture the fact that ambiguous continuation value reduces the utility of the decision maker since ambiguity aversion implies that the representative household dislikes mean-preserving spread in expected continuation value due to the existence of model uncertainty. Behind recursive smooth model of ambiguity, there is the notion of consequentialism and dynamic consistency, who is the most intuitive way when we talk about decision makings within the lens of business cycles. And it admits a tractable Bellman equation formulation.

Denote value function as $J_t \equiv J\left(\left\{K_{j,t}\right\}, a_{t-1}, z_t, \zeta_t, \psi_t\right)$. Standard dynamic programming argument can be applied resulting the following Bellman equation for the representative household at stage 2:

$$J_t = \max_{C_t, \left\{K_{j,t+1}\right\}} u\left(C_t\right) + \beta\phi^{-1}\left(\int_{\mathcal{R}} \phi\left(E_{t,2}^{\omega_{t+1}}\left[J_{t+1}\right]\right) f_t\left(\omega_{t+1}\right) d\omega_{t+1}\right) \tag{2.31}$$

subject to

$$P_t C_t + P_t \int_J I_{j,t} dj = \int_J W_{j,t} N_{j,t} didj + \int_J R_{j,t} K_{j,t} dj + \int_J \Pi_{j,t} dj \tag{2.32}$$

and

$$I_{j,t} = K_{j,t+1} - (1-\delta) K_{j,t} \tag{2.33}$$

Here $E_{t,2}^{\omega_{t+1}}[\cdot]$ stands for the mathematical expectation conditioned on $\mathcal{I}_{t,2}$ under a particular model ω_{t+1} for cross-sectional mean of tomorrow's idiosyncratic productivity shocks. And $f_t\left(\omega_{t+1}\right)$ stands for the probability density function for households' period t prior over tomorrow's ambiguity ω_{t+1}. Since period t knowledge does not reveal any information over tomorrow's ambiguity, prior belief over tomorrow's ambiguity ω_{t+1} at stage 2 coincides with that at stage 0.

Preference of the representative household at stage 1. Similar to the simple model, at stage 1, preference of the representative household is given by the smooth model of ambiguity

$$\int_{\Omega_t} \phi \left(E^{\omega_t}_{j,t,1} \left[J \left(\{ K_{j,t} \}, a_{t-1}, z_t, \zeta_t, \psi_t \right) \right] \right) \widetilde{f}^h_{j,t,1} \left(\omega_t \right) d\omega_t$$

where the interim belief follows the smooth rule of updating

$$\widetilde{f}^h_{j,t,1} \left(\omega_t \right) \propto \underbrace{\frac{\phi' \left(E^{\omega_t}_{j,t,0} \left[J \left(\{ K_{j,t} \}, a_{t-1}, z_t, \zeta_t, \psi_t \right) \right] \right)}{\phi' \left(E^{\omega_t}_{j,t,1} \left[J \left(\{ K_{j,t} \}, a_{t-1}, z_t, \zeta_t, \psi_t \right) \right] \right)}}_{\text{Weights}} \underbrace{f \left(x_{j,t} | \omega_t \right) f_t \left(\omega_t \right)}_{\text{Bayesian Kernel}}$$

Firm problem at stage 1. Firm problem is formulated in a similar fashion as the simple model:

$$\int_{\Omega_t} \phi \left(E^{\omega_t}_{j,t,1} \left[\frac{U' \left(C_t \right)}{P_t} \left(P_{j,t} Y_{j,t} - W_{j,t} N_{j,t} - R_{j,t} K_{j,t} \right) \right] \right) \widetilde{f}^f_{j,t,1} \left(\omega_t \right) d\omega_t$$

where the interim belief system satisfying the extended smooth rule of updating to ensure dynamic consistency:

$$\widetilde{f}^w_{j,t,1} \left(\omega_t \right) \propto \underbrace{\frac{\phi' \left(E^{\omega_t}_{j,t,0} \left[J \left(\{ K_{j,t} \}, a_{t-1}, z_t, \zeta_t, \psi_t \right) \right] \right)}{\phi' \left(E^{\omega_t}_{j,t,1} \left[\ln \left(C_t \right) - \chi \int_J \frac{N^{1+\epsilon}_{j,t}}{1+\epsilon} dj \right] \right)}}_{\text{Weights}} \underbrace{f \left(x_{j,t} | \omega_t \right) f_t \left(\omega_t \right)}_{\text{Bayesian Kernel}}$$

To close up the description of the model, we make the following assumptions on functional forms

Assumption 2.4. (Log-Exponential) We assume $u \left(C_t \right) = \ln C_t$ and $\phi \left(x \right) = -\frac{1}{\lambda} e^{-\lambda x}$ with $\lambda \geq 0$.

In what follows, we leave out the optimality conditions to Appendix. And directly move to the discussion over our quantitative methodology in the approximation of conditional log-normal equilibrium, which is closely related to what we have done in Section 2.4.3.

2.6.2. Quantitative Methodology

The key feature of the smooth model of ambiguity is that decision makers would invoke a distorted (relative to Bayesian posterior) posterior belief over the set of the possible model when evaluating marginal effects of factors supply and demand. In our RBC extension, the belief distortion at stage 1 is given by

$$M_{t,1}\left(\omega_t\right) = e^{-\lambda E^{\omega_t}_{t,0}[J_t]} \tag{2.34}$$

and the belief distortion at stage 2 is given by

$$M_{t,2}\left(\omega_{t+1}\right) = e^{-\lambda E^{\omega_{t+1}}_{t,2}[J_{t+1}]} \tag{2.35}$$

To quantitatively pin down the whole equilibrium, expected value functions as functions of concurrent and tomorrow's ambiguity, $\left\{E^{\omega_t}_{t,0}\left[J_t\right], E^{\omega_{t+1}}_{t,2}\left[J_{t+1}\right]\right\}$, have to be approximated jointly with policy rules.

For any variable S_t of interest, we use hatted-lower-case $\widehat{s}_t$ to denote the log-deviation from its ambiguous steady state (A-SS):

$$\widehat{s}_t \equiv \ln\left(S_t\right) - \ln\left(S^*\right) - \overline{h}_s\left(\overline{\psi}\right)$$

where S^* stands for the deterministic steady state (D-SS) level of S_t and $\overline{h}_s\left(\overline{\psi}\right)$ is a function of $\overline{\psi}$, which takes into account the impacts of ambiguity on S_t at A-SS.

Focusing on conditional Log-Normal equilibrium (Definition 2.2), we propose the following policy rules for island employment $\widehat{n}_{j,t}$, output $\widehat{y}_{j,t}$, wage rate $\widehat{w}_{j,t}$ and rental rental rate of capital $\widehat{r}_{j,t}$ at stage 1 of period t:

$$\begin{aligned}
\widehat{y}_{j,t} &= \kappa_{yk}\widehat{k}_t + \kappa_{ya}a_{t-1} + \kappa_{yx}\left(\psi_t\right)x_{j,t} + \widehat{h}_y\left(\psi_t\right)\\
\widehat{n}_{j,t} &= \kappa_{nk}\widehat{k}_t + \kappa_{na}a_{t-1} + \kappa_{nx}\left(\psi_t\right)x_{j,t} + \widehat{h}_n\left(\psi_t\right)\\
\widehat{w}_{j,t} &= \kappa_{wk}\widehat{k}_t + \kappa_{wa}a_{t-1} + \kappa_{wx}\left(\psi_t\right)x_{j,t} + \widehat{h}_w\left(\psi_t\right)\\
\widehat{r}_{j,t} &= \kappa_{rk}\widehat{k}_t + \kappa_{ra}a_{t-1} + \kappa_{rx}\left(\psi_t\right)x_{j,t} + \widehat{h}_r\left(\psi_t\right)
\end{aligned}$$

and the following policy rules for consumption $\widehat{c}_t$, investment $\widehat{i}_t$ and capital stock tomorrow $\widehat{k}_{t+1}$ at stage 2 of period t:

$$\begin{aligned}
\widehat{c}_t &= \kappa_{ck}\widehat{k}_t + \kappa_{ca}a_{t-1} + \kappa_{cz}\left(\psi_t\right)z_t + \kappa_{c\zeta}\left(\psi_t\right)\zeta_t + \widehat{h}_c\left(\psi_t\right)\\
\widehat{i}_t &= \kappa_{ik}\widehat{k}_t + \kappa_{ia}a_{t-1} + \kappa_{iz}\left(\psi_t\right)z_t + \kappa_{i\zeta}\left(\psi_t\right)\zeta_t + \widehat{h}_i\left(\psi_t\right)\\
\widehat{k}_{t+1} &= \kappa_{kk}\widehat{k}_t + \kappa_{ka}a_{t-1} + \kappa_{kz}\left(\psi_t\right)z_t + \kappa_{k\zeta}\left(\psi_t\right)\zeta_t + \widehat{h}_k\left(\psi_t\right)
\end{aligned}$$

To elaborate a bit, first of all, the log-deviations of variables of interest are assumed to be from the A-SS instead of the D-SS. This is because the amount of ambiguity at A-SS $\overline{\psi}$ has a non-negligible first-moment impact on allocations when decision makers are ambiguity averse. Secondly, stage 1 variables are measurable with respect to stage 1 information sets. Since idiosyncratic productivity shocks are i.i.d across time and across islands, past information can be effectively summarized into $\left\{\widehat{k}_t, a_{t-1}, \psi_t\right\}$. Therefore stage 1 variables, i.e., $\left\{\widehat{y}_{j,t}, \widehat{n}_{j,t}, \widehat{w}_{j,t}, \widehat{r}_{j,t}\right\}$, are functions of $\left\{\widehat{k}_t, a_{t-1}\psi_t, x_{j,t},\right\}$ only. Similar arguments apply to stage 2 variables, i.e., $\left\{\widehat{c}_t, \widehat{i}_t, \widehat{k}_{t+1}\right\}$. Thirdly, fixing the amount of ambiguity ψ_t, the policy rules are linear in productivity shocks either the aggregate or the idiosyncratic. This corresponds to the standard log-linearization when there is no ambiguity shock at all, i.e. $\psi_t = \overline{\psi}$ for $\forall t$. When there presents ambiguity shock, we allow it to interact with productivity shocks $x_{j,t}$, z_t and ζ_t in a possibly non-linear way. This reflects the fact that ambiguity shock is capable of generating time-varying response to productivity shocks, i.e. $\left\{\kappa_{*x}(\psi_t), \kappa_{*z}(\psi_t), \kappa_{*\zeta}(\psi_t)\right\}$. In addition, if decision makers are ambiguity averse, ambiguity shock has the first moment impacts manifested by the possibly non-linear functions $\widehat{h}_*(\psi_t)$. This reflects fluctuations in degree of pessimism over the short-run outlooks of the economy. In sum, the proposed policy rules can be understood as a semi-linear perturbation around the ambiguous steady state where we allow for interactions between productivity and ambiguity shocks as well as all higher-order terms of ambiguity shocks.

To approximate the conditional Log-Normal equilibrium with the proposed policy rules, we first implement a quadratic approximation of the value function

$$
\begin{aligned}
J_t =& J^* + \overline{h}_J + \kappa_{Jk}\widehat{k}_t + \kappa_{Ja}a_{t-1} + \kappa_{Jz,t}z_t + \kappa_{J\zeta,t}\zeta_t \\
&+ \kappa_{Jka}\widehat{k}_t a_{t-1} + \kappa_{Jkz,t}\widehat{k}_t z_t + \kappa_{Jk\zeta,t}\widehat{k}_t\zeta_t + \kappa_{Jaz,t}a_{t-1}z_t + \kappa_{Ja\zeta,t}a_{t-1}\zeta_t + \kappa_{Jz\zeta}z_t\zeta_t \\
&+ \frac{1}{2}\kappa_{Jkk}\widehat{k}_t^2 + \frac{1}{2}\kappa_{Jaa}a_{t-1}^2 + \frac{1}{2}\kappa_{Jzz,t}z_t^2 + \frac{1}{2}\kappa_{J\zeta\zeta,t}\zeta_t^2 + \widehat{h}_J(\psi_t)
\end{aligned}
$$

Plug this back to the value function recursion (2.31), we can approximate expected value functions by

$$
E_{t,0}^{\omega_t}[J_t] \approx constant_t + \kappa_{Jz,t}(\psi_t)\,\omega_t + \frac{1}{2}\kappa_{Jzz,t}(\psi_t)\,\omega_t^2
$$

and

$$
\begin{aligned}
E_{t,2}^{\omega_{t+1}}[\beta J_{t+1}] \approx constant_t &+ \beta\left(\int_{\mathcal{R}}\kappa_{Jz,t}(\psi_{t+1})\,dF(\psi_{t+1}|\psi_t)\right)\omega_{t+1} \\
&+ \frac{1}{2}\left(\int_{\mathcal{R}}\kappa_{Jz,t}(\psi_{t+1})\,dF(\psi_{t+1}|\psi_t)\right)\omega_{t+1}^2
\end{aligned}
$$

where $\kappa_{Jz,t}$ and $\kappa_{Jzz,t}$ are functions of undetermined coefficients in the above proposed policy rules. A direct implication of the quadratic forms in belief distortions $M_{t,1}(\omega_t)$ and $M_{t,2}(\omega_{t+1})$ is Normality in posterior belief over the set of possible models both at stage 1 and stage 2.

In the next step, we log-linearize the optimality conditions derived in Appendix around the A-SS. While doing this, we seriously take into account that at stage 1 and 2 posterior belief over cross-sectional mean of idiosyncratic productivity shocks are distorted in a way consistent with the estimated belief distortions $M_{t,1}(\omega_t)$ and $M_{t,2}(\omega_{t+1})$. By doing this, we share the same spirit with Ilut and Schneider (2014) and Ilut and Saijo (2016) in dealing with log-linearization with distorted subjective belief. Plugging the proposed policy rules into the log-linearized optimality conditions, we arrive at a large system of undetermined coefficients. To note that in the practical implementation, we discretize the AR(1) process of ambiguity to transform functions over ψ_t into a finite number of coefficients, each of which corresponds to the value of the function at a specific value for ψ_t. Such a huge system of undetermined coefficients can be solved quantitatively with the restriction that $\kappa_{kk} < 1$ to ensure TVC is not violated. Detailed math can be found in the Appendix.

2.6.3. Calibration

Table 2.1 summarizes the parameters used in the calibrated version of our baseline model. Discount factor β is 0.99; Frisch elasticity of labor supply is 2; capital share in production is 0.36, and depreciation rate of capital is 0.025. θ is chosen to be 1 corresponding to Cobb-Douglas aggregation technology over island commodities, which implies $y_t = \int_J y_{j,t} dj$. Finally, χ is chosen to be 4.47 to ensure that $1/3$ of time is devoted to working in the deterministic steady state.

The persistence of aggregate productivity shock ρ is chosen to be 0.95, a conventional value in the literature. Following Angeletos, Collard, and Dellas (2016b), the persistence of ambiguity shock ρ_ψ is 0.75 indicating a 2.5 quarters half-life of ambiguity shock, which resembles aggregate demand shock in Blanchard and Quah (1989). It remains to specify the standard deviations of the three exogenous shocks: the aggregate productivity shock σ_ζ, the idiosyncratic productivity shock σ_ι and the ambiguity shock σ_τ, as well as the amount of ambiguity at A-SS $\overline{\psi}$ and the degree of ambiguity aversion λ.

Following Angeletos, Collard, and Dellas (2016b), we select these 5 parameters to minimize the distance between the model implied standard deviations

Table 2.1. *Model Parameters*

Parameters	Role	Value
β	discount factor	0.99
ϵ	inverse Frisch elasticity	0.5
α	capital share	0.36
δ	depreciation rate	0.025
θ	Cobb-Douglas aggregation	1
χ	1/3 hours at D-SS	4.47
ρ	persistence of agg. productivity shock	0.95
ρ_ψ	persistence of ambiguity shock	0.75
$100\sigma_\zeta$	std. dev. of agg. productivity shock	0.78
$100\sigma_\iota$	std. dev. of island productivity shock	9.0
σ_τ	std. dev. of ambiguity shock	0.47
$\overline{\psi}$	amount of ambiguity $e^{\overline{\psi}}$ at A-SS	-5.05
λ	Degree of ambiguity aversion	12.5

of output, consumption, hours, investment and labor productivity and their data counterparts, where the distances of each variable are weighted by the precision of model-based estimators. We arrive at the calibration such that $\sigma_\zeta = 0.0078$ and $\sigma_\iota = 0.090$. The process of the ambiguity shock has a long-run mean $e^{\overline{\psi}} = 0.0064$ and standard deviation $\sigma_\tau = 0.47$. Finally, the degree of ambiguity aversion is 12.5.[17]

2.6.4. Business-Cycle Moments and Aggregate Co-movements

Business-cycle moments. Table 2.2 summarizes key moments of aggregate variables in the US data over the period of 1971Q1-2014Q4 (column 1) and in our calibrated baseline model (column 2). The overall empirical fit of the calibrated baseline model is quite well.

At the core of such overall fit are the balancing roles between the two aggregate shocks, i.e., aggregate productivity shock ζ_t and ambiguity shock ψ_t. Column 3 and 4 report the business cycle moments for aggregate variables when there are only aggregate productivity shocks by setting $\sigma_\tau = 0$ or am-

[17]To note that, even though there are 5 parameters to match 5 moments, we cannot ensure perfect matching. This is because we are not doing the unconstrained minimization. We are minimizing the objective under the constraint that $\lambda e^{\overline{\psi}}$ is bounded above. Such a constraint implicitly assumes there should not be too much deviation in belief from objective one at A-SS and also ensures the uniqueness of A-SS.

Table 2.2. *Bandpass-Filtered Moments of Aggregate Variables*

	Data(1971Q1-2014Q4)	Baseline Model	A Only	ψ Only
Standard Deviations				
$stddev(y)$	1.45	1.72	1.28	1.17
$stddev(c)$	0.87	0.72	0.38	0.63
$stddev(n)$	1.76	1.93	0.60	1.83
$stddev(i)$	5.45	4.51	3.75	2.63
$stddev(y/n)$	0.84	0.96	0.69	0.67
Correlations				
$corr(c,y)$	0.88	0.92	0.92	0.98
$corr(n,y)$	0.88	0.87	0.99	0.99
$corr(i,y)$	0.95	0.99	0.99	0.99
$corr(c,n)$	0.85	0.93	0.90	0.97
$corr(c,i)$	0.80	0.84	0.88	0.95
$corr(i,n)$	0.85	0.80	0.99	0.99
Corr. with Productivity				
$corr(y,y/n)$	-0.11	0.05	0.99	-0.99
$corr(n,y/n)$	-0.57	-0.45	0.97	-0.99

Note: The first column reports moments for US data from 1971Q1 to 2014Q4. The second column reports moments in our baseline model. Column 3 and 4 report moments generated by some models when there are only aggregate productivity shocks or ambiguity shocks respectively. All moments are band-pass filtered at frequencies 6-32 quarters.

biguity shocks by setting $\sigma_\zeta = 0$ respectively. As in the case of standard RBC, when there are only aggregate productivity shocks, the model fails to generate enough fluctuations in hours and predicts counterfactually high positive correlations between output (or hours) and labor productivity. On the contrary, when there are only ambiguity shocks, the model generates too much volatility in hours and predicts almost perfectly negative correlations between output or hours and labor productivity. In combination, the overall fit is achieved through a natural balancing by our calibrated baseline model.

Aggregate co-movements. Impulse response functions of key aggregate variables to a positive ambiguity shock are reported in . The aggregate co-movement patterns are akin to those of the confidence shocks in Angeletos, Collard, and Dellas (2016b), Huo and Takayama (2015) and Ilut and Saijo (2016), where a positive ambiguity shock generates a drop of aggregate quantities, i.e., output, consumption, hours and investment, while, at the same time, an increase in labor productivity in a way consistent with interpretation of aggregate demand shocks.

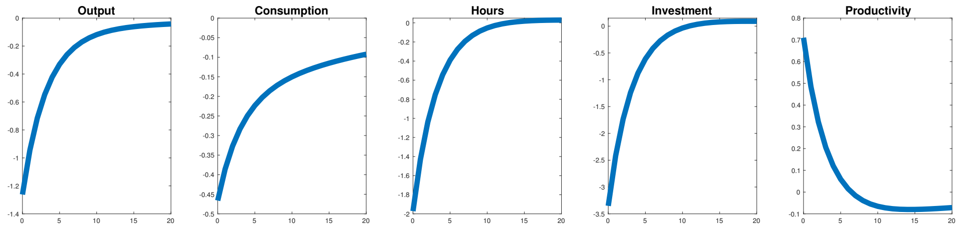


Figure 2.2. *Impulse responses to one standard deviation of positive ambiguity shock*

What drives the co-movements pattern behind the these IRFs is the fluctuations of in the degree of the pessimism over the short-run outlooks of the economy. By construction, a positive ambiguity shock deepens the degree of the pessimism of all agents over the cross-sectional mean of idiosyncratic productivity shocks for all periods onwards. From the perspective of the firms, such an increased pessimism means into a depressed expectation over the demand of own commodities. In response, firms reduce demand for labor and capital, generating downward pressure on factors prices. From the perspective of the households, this implies a modest decrease in expected permanent income. In response, consumption drops. At the same time, the modest drop in expected permanent income, unlike the case for aggregate productivity shocks, restricts the strength of wealth effect. Given the fact that households understand the drop in factors prices only last for the near future, hours and investment decrease in equilibrium since the relevant substitution effect dominates the opposing wealth effect. In sum, ambiguity shocks generate aggregate comovements patterns depicted in Figure 2.2.

Labor wedges. Our calibrated baseline model does a considerably good job in capturing features of data on hours. We interpret such a goodness of fit through the lens of labor wedge analysis in line with Chari, Kehoe, and McGrattan (2007). The key idea is to interpret aggregate data on quantities and prices as wedges in optimality conditions of a text-book RBC model. Then the question of interest is the following: how will the data generated by our calibrated baseline model be translated into wedges, especially labor wedges? It turns out that ambiguity shocks generate empirical relevant countercyclical labor wedges. To economize space, detailed math regarding the calculation of labor wedge are moved to the Appendix.

Table 2.3 compares moments of labor wedge estimated from US data over the period of 1971Q1-2014Q4 with their model counterparts. It turns out our calibrated baseline model does a considerable good job in capturing cyclical behaviors in labor wedge. It is the exactly the aggregate demand shock na-

Table 2.3. *Bandpass-filtered Moments: Labor Wedges*

	Data$_{(1971Q1\text{-}2014Q4)}$	Baseline Model
Stddev	2.17	2.20
Correlation with y	-0.75	-0.66

Note: The first column reports moments of labor wedge for US data from 1971Q1 to 2014Q4. The second column reports corresponding moments in our baseline model. All moments are band-pass filtered at frequencies 6-32 quarters. Details can be found in Appendix.

ture embedded in ambiguity shock that helps us in explaining labor market dynamics to a large extent. A positive ambiguity shock makes the households more pessimistic over current economic condition hence expecting consumption to be lower. This drives down the expected marginal rate of intra-temporal substitution between leisure and consumption, providing more incentives for households to supply labor. Therefore, a positive ambiguity shock acts as a subsidy on labor supply. On the contrary, a positive ambiguity shock makes island j firms believe demand conditions are turning bad and therefore generates downward pressure on labor demand. In this sense, a positive ambiguity shock acts as a tax on labor demand. In combine, firm side effect dominates household side effect. Therefore, a positive ambiguity shock act as a total labor wedge on island j. This implies that, on average, a positive ambiguity shock generates larger labor wedge hence decreases hours in equilibrium.

2.6.5. Belief Divergence

Following Assumption 2.3, professional forecasters in the model are assumed to be ambiguity neutral. Also, to capture the fact that these professional forecasters have better information position than private agents inside the economy do, we provide with the professional forecasters in our model with one additional private information over average productivity $\int_J x_{j,t}dj$:

$$s_{j,t} = \int_J x_{j,t}dj + \xi_{j,t} \quad \text{with} \quad \xi_{j,t} \sim N\left(0, \sigma_\xi^2\right)$$

We calibrate the standard deviation of this additional private information σ_ξ^2 to match the standard deviation of the cross-sectional dispersion in SPF data over the period of 1987Q1-2014Q4 resulting in $100\sigma_\xi = 0.52$.

Table 2.4 reports the standard deviation and correlation with output y for the US data. Over the period of 1987Q1-2014Q4, the starting point of time of which corresponds to the beginning of great moderation, there displays little variation in belief divergence such that standard deviation is only 0.05 percentage points. However, it is significantly negative correlated with output (-0.39). Our calibrated baseline model captures exactly this pattern. To note that none of our parameters are chosen to match the correlation with output in Table 2.4. And the cyclical pattern of this moment is solely due to ambiguity shock in a way such that a positive ambiguity is predicted to drive up belief divergence. Therefore, we can conclude that ambiguity shocks capture salient features of data in belief divergence pretty well.

Table 2.4. *Bandpass-Filtered Moments of Belief Divergence*

	Data(1987Q1-2014Q4)	Baseline Model
Stddev	0.05	0.05
Correlation with y	-0.39	-0.67

Note: The first column reports moments of belief divergence for US data from 1987Q1 to 2014Q4. The second column reports corresponding moments in our baseline model. All moments are band-pass filtered at frequencies 6-32 quarters.

2.6.6. Estimated Market Confidence v.s. Sentiment Index

To further validate our theory, we address the following question: can our calibrated baseline model replicate the movements in market confidence as proxied by Sentiment Index in the data?

Figure 2.3 compares the estimated time-series for market confidence, hours, belief divergence and output together with their data counterparts between 1987Q1-2014Q4. To construct these estimated time-series, we first select the sequence of ambiguity and aggregate TFP shocks that can perfectly back-out output and the cross-sectional dispersion of output forecast in SPF. When we backout these shocks we assume that the economy initially stays at the ambiguous steady state.[18] We then construct a sequence of simulated market confidence following Definition 2.3. Notably, the simulated market confidence ψ_t closely tracks Sentiment Index in Michigan Survey of Consumers, especially for the periods corresponding to NBER recessions. Recall that we do not use any information on Sentiment Index in the construction of estimated time-series for ambiguity and aggregate TFP shocks. Such an empirical fit provides additional

[18]These initial conditions are consistent with data that 1985 is neither a boom or recession. Since the first 8 observations will be dropped after band-pass filtering, our findings are actually quite robust to the choice of initial conditions.

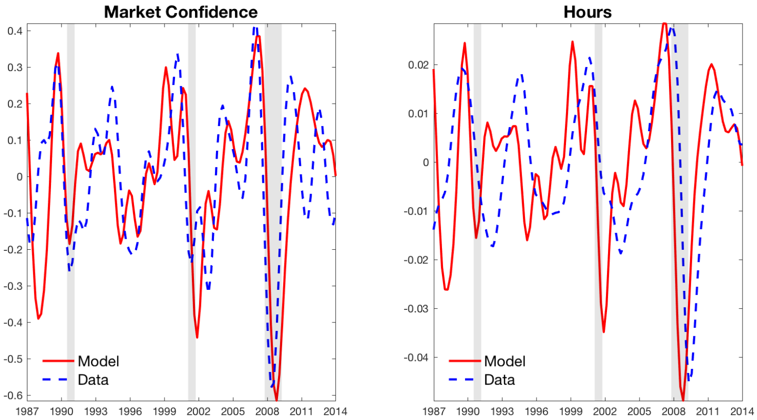


Figure 2.3. Estimated Times-Series vs Empirical Proxies

Note: The figure plots time-series for market confidence and hours estimated from our model and their data counterparts between 1987Q1-2014Q4. In the data, market confidence is proxied by Consumer Sentiment Index from Michigan Survey of Consumer. All moments are band-pass filtered at frequencies 6-32 quarters.

validation to our theory. Also note that, simulated time-series of hours also closely tracks its data counterpart.

2.7. Conclusion

We develop a novel theory of ambiguity-driven business cycles that contributes to explain the co-movements across market confidence, belief divergence, and the aggregate economy. We work on a standard RBC model with multiple firms and multiple commodities and further extend it with ambiguity averse preference represented by the smooth model of ambiguity. Within the smooth model of ambiguity, we contribute to the literature with a Bayesian formulation of ambiguity shock, namely shock to the variance of agents' prior belief over possible models.

Within a simple model without capital, we demonstrate that a positive ambiguity shock makes all agents, who are ambiguity averse, behave as if they believe the aggregate fundamental is turning bad and becoming more volatile. Such dual impacts of ambiguity shock generate endogenous movements across confidence and uncertainty. When the economy features imperfect coordination due to incomplete information, dual impacts of a positive ambiguity shock translates into depressed belief over aggregate demand and the increased incentives to use private information both when making output decisions or output forecasts. The former maps into depressed market confidence and the latter maps into heightened belief divergence. And finally, aggregate output falls due to the increase in the economy-wide pessimism over aggregate demand. In this sense, ambiguity shock in our paper is nothing more than a particular formulation of the aggregate demand shock. In combination, a positive ambiguity shock generates recession with depressed market confidence and heightened belief divergence.

We further explore the quantitative potential of our theory within a dynamic RBC model. Ambiguity shock is shown to be capable of generating co-movements across real quantities together with counter-cyclical labor productivity and labor wedge. Our model is also capable of capturing cyclicalities in belief divergence as measured by the cross-sectional dispersion in output forecast in SPF dataset. Also, the estimated time series of market confidence closely tracks Sentiment Index in Michigan Survey of Consumer. Therefore, we conclude that fluctuations in market confidence, belief divergence, and the aggregate economy are nothing more than the many shades of ambiguity shock, not only qualitatively but also quantitatively.

Appendix of Chapter 2

2.A. Derivations and Proofs

Derivation of Equation (2.15). FOC for the island j workers' problem is such that

$$\int_{\mathcal{R}} \phi' \left(E^{\omega_t}_{j,t,1} \left[\frac{C_t^{1-\gamma} - 1}{1-\gamma} - \chi \int_J \frac{N_{j,t}^{1+\epsilon}}{1+\epsilon} dj \right] \right) E^{\omega_t}_{j,t,1} \left[C_t^{-\gamma} W_{j,t} - \chi N_{j,t}^{\epsilon} \right] \widetilde{f}^h_{j,t,1} (\omega_t) \, d\omega_t = 0$$

Plugging in the expression for $\widetilde{f}^h_{j,t,1}(\omega_t)$ given by (2.6), we arrive at (2.12) where distorted posterior belief over the set of possible models can be shown given by (2.14). Similar procedures lead to (2.13). Then in the last step, combining (2.12) and (2.13) together leads to (2.15). ■

Proof of Lemma 2.1. Under conditional log-normal equilibrium, we have that

$$\begin{aligned} y_{j,t} =& y^* + \overline{h}_y + \kappa_{yx,t} x_{j,t} + \widehat{h}_y (\psi_t) \\ n_{j,t} =& n^* + \overline{h}_n + \kappa_{nx,t} x_{j,t} + \widehat{h}_n (\psi_t) \\ y_t =& \int_J y_{j,t} dj + \frac{1}{2} \left(1 - \frac{1}{\theta} \right) d_{y,j}^2 \end{aligned}$$

where $d_{y,j} \equiv \kappa_{yx,t}^2 \sigma_t^2$ denotes the cross-sectional dispersion in island outputs. We ignore $d_{y,j}$ in the approximation without loss of generality since they are of second order impacts at the aggregates and have no impacts at all on the cyclical behaviours of belief divergence.

Define $\overline{S} = e^{s^* + \overline{h}_s}$ for any variable of interest S. Quadratic approximation over period utility of the representative household is given by

$$\begin{aligned} & \frac{Y_t^{1-\gamma} - 1}{1-\gamma} - \chi \int_J \frac{N_{j,t}^{1+\epsilon}}{1+\epsilon} dj \\ \approx & \frac{\overline{Y}^{1-\gamma}}{1-\gamma} \left[1 + (1-\gamma) \widehat{y}_t + \frac{1}{2} (1-\gamma)^2 \widehat{y}_t^2 \right] - \frac{1}{1-\gamma} - \chi \frac{\overline{N}^{1+\epsilon}}{1+\epsilon} \int_J \left(1 + (1+\epsilon) \widehat{n}_{j,t} + \frac{1}{2} (1+\epsilon)^2 \widehat{n}_{j,t}^2 \right) dj \\ =& Const. + \overline{Y}^{1-\gamma} \widehat{y}_t + \frac{1}{2} (1-\gamma) \overline{Y}^{1-\gamma} \widehat{y}_t - \int_J \left(\chi \overline{N}^{1+\epsilon} \widehat{n}_{j,t} + \frac{1}{2} (1+\epsilon) \chi \overline{N}^{1+\epsilon} \widehat{n}_{j,t}^2 \right) dj \end{aligned}$$

Further define ex-ante (stage 0) expected utility given a particular model ω_t as

$\overline{J}_t(\omega_t)$ such that

$$\overline{J}_t(\omega_t) \equiv E^{\omega_t}_{t,0}\left[\frac{Y_t^{1-\gamma}-1}{1-\gamma} - \chi\int_J \frac{N_{j,t}^{1+\epsilon}}{1+\epsilon}dj\right]$$

It turns out $\overline{J}_t(\omega_t)$ can be quadratically approximated by

$$\overline{J}_t(\omega_t) \approx Const_t + \kappa_{J\omega}(\psi_t,\lambda)\,\omega_t + \frac{1}{2}\kappa_{J\omega\omega}(\psi_t,\lambda)\,\omega_t^2$$

where coefficients of linear and quadratic terms are given by

$$\begin{aligned}
\kappa_{J\omega}(\psi_t,\lambda) =& (Y^*)^{1-\gamma}\left(1+(1-\gamma)\overline{h}_y\right)\kappa_{ya_j}(\psi_t,\lambda) - \chi(N^*)^{1+\epsilon}\left(1+(1+\epsilon)\overline{h}_n\right)\kappa_{na_j}(\psi_t,\lambda)\\
\kappa_{J\omega\omega}(\psi_t,\lambda) =& (Y^*)^{1-\gamma}\left(1+(1-\gamma)\overline{h}_y\right)(1-\gamma)\,\kappa^2_{ya_jt}(\psi_t,\lambda)\\
&- \chi(N^*)^{1+\epsilon}\left(1+(1+\epsilon)\overline{h}_n\right)(1+\epsilon)\,\kappa^2_{na_j}(\psi_t,\lambda)
\end{aligned}$$

To derive such an approximation, we first approximate $e^{(1-\gamma)\overline{h}_y}$ and $e^{(1+\epsilon)\overline{h}_n}$ by $1+(1-\gamma)\overline{h}_n$ and $1+(1+\epsilon)\overline{h}_n$ respectively. This is doable in the macro application since we want to restrict the A-SS impact of ambiguity to avoid too much statistical sophistication. Then we ignore the intersection terms between ω_t and $x_{j,t}$ or between ω_t and $\widehat{h}_s(\psi_t)$ with $s \in \{y,n\}$. This is doable since those terms are negligible comparing to $\kappa_{J\omega}\omega_t$.

Finally, quadratic approximation over $\overline{J}_t(\omega_t)$ implies that belief distortion in (2.17) is of exponential quadratic form. Therefore, posterior belief over the set of possible models $\widetilde{f}_{j,t,1}(\omega_t)$ is Normal since the kernel would also be quadratic in ω_t. This leads to a Normal density with the mean μ_t and variance σ_t^2 given by

$$\mu_t = \left(\frac{e^{\psi_t}+g_\sigma(\psi_t,\lambda)}{\sigma_\zeta^2+\sigma_\iota^2+e^{\psi_t}+g_\sigma(\psi_t,\lambda)}\right)x_{j,t} + \left(\frac{\sigma_\zeta^2+\sigma_\iota^2}{\sigma_\zeta^2+\sigma_\iota^2+e^{\psi_t}+g_\sigma(\psi_t,\lambda)}\right)g_\mu(\psi_t,\lambda) \tag{2.36}$$

and

$$\sigma_t^2 = \left(\frac{\sigma_\zeta^2+\sigma_\iota^2}{\sigma_\zeta^2+\sigma_\iota^2+e^{\psi_t}+g_\sigma(\psi_t,\lambda)}\right)\left(e^{\psi_t}+g_\sigma(\psi_t,\lambda)\right) \tag{2.37}$$

where the distortions in mean $g_\mu(\psi_t, \lambda)$ and in variance $g_\sigma(\psi_t, \lambda)$ are given by

$$
\begin{aligned}
g_\mu(\psi_t, \lambda) &= -\lambda \kappa_{J\omega}(\psi_t, \lambda)\left(\frac{e^{\psi_t}}{1 + \lambda \kappa_{J\omega\omega}(\psi_t, \lambda) e^{\psi_t}}\right) \\
g_\sigma(\psi_t, \lambda) &= -\left(\frac{\lambda \kappa_{J\omega\omega}(\psi_t, \lambda) e^{\psi_t}}{1 + \lambda \kappa_{J\omega\omega}(\psi_t, \lambda) e^{\psi_t}}\right) e^{\psi_t}
\end{aligned}
$$

In any well-defined equilibrium, it has to be the case that $e^{\psi_t} + g_\sigma(\psi_t, \lambda) > 0$. Otherwise, there would be no well defined distroted posterior belief $\widetilde{f}_{j,t,1}(\omega_t)$ in the sense that its kernel density would be explosive. ∎

Proof of Proposition 2.1. Suppose that the conditional Log-Normal equilibrium takes the following form:

$$
\begin{aligned}
y_{j,t} \equiv \ln Y_{j,t} =& y^* + \overline{h}_y(\overline{\psi}) + \kappa_{ya_j}(\psi_t, \lambda) \cdot a_{j,t} + \widehat{h}_y(\psi_t, \lambda) \\
n_{j,t} \equiv \ln N_{j,t} =& n^* + \overline{h}_n(\overline{\psi}) + \kappa_{na_j}(\psi_t, \lambda) \cdot a_{j,t} + \widehat{h}_n(\psi_t, \lambda) \\
y_t \equiv \ln Y_t =& y^* + \overline{h}_y(\overline{\psi}) + \kappa_{ya_j}(\psi_t, \lambda) \cdot \int_J a_{j,t} dj + \widehat{h}_y(\psi_t, \lambda)
\end{aligned}
$$

where we ignore dispersion adjustment of aggregate output in the approximation without loss of generality since they are of second order impacts at the aggregates and have no impacts at all on the cyclical behaviours of belief divergence. Then at D-SS, we have the following

$$
\ln(\chi) + (1 + \epsilon) n^* = \ln(1 - \alpha) + (1 - \gamma) y^*
$$

While, at A-SS, impacts of ambiguity shocks at A-SS denoted by $\overline{h}_s$ $s \in \{n, y\}$ must satisfy the following

$$
\overline{h}_n = (1 - \gamma)\overline{h}_y + \left(\frac{1}{\theta} - \gamma\right) H_y(\overline{\psi}, \lambda) \tag{2.38}
$$

Here $H_y(\overline{\psi}, \lambda)$ denotes degree of pessimism of island j agents over aggregate output y_t at the A-SS. Under the proposed conditional Log-Normal equilibrium, it is given by

$$
H_y(\overline{\psi}, \lambda) = \kappa_{ya_j}(\overline{\psi}, \lambda)\left(\int_{\mathcal{R}} E^{\omega_t}_{j,t,1}\left[\int_J a_{j,t} dj\right] \widetilde{f}_{j,t,1}(\omega_t) d\omega_t\right)\Bigg|_{x_{j,t}=0, \psi_t=\overline{\psi}}
$$

To understand why this is the case, recall that ambiguous steady state refers to the state the economy converges to (a) in the absence of any shocks, i.e., $a_{j,t} = 0$, but (b) taking into account of the existence of ambiguity, i.e., evaluating

$\int_{\mathcal{R}} E^{\omega_t}_{j,t,1}\left[\int_J a_{j,t}dj\right]\widetilde{f}_{j,t,1}(\omega_t)\,d\omega_t$ at $\psi_t = \overline{\psi} \neq -\infty$. Alternatively, we can interpret $H_y(\overline{\psi})$ from the perspective of distorted subjective beliefs of all agents. At A-SS, the amount of ambiguity $\overline{\psi}$ plays a non-trivial role in the sense that even at A-SS, agent's subjective belief over average productivity is distorted in the mean. This mean distortion has to be respected when we evaluate the A-SS, leading to a non-zero term $H_y(\overline{\psi})$. Similar arguments can be found in Ilut and Schneider (2014) and Ilut and Saijo (2016) in the context of "worst case" belief due to multiple prior preferences. Following (2.36), we have that

$$H_y(\overline{\psi},\lambda) = \kappa_{ya_j}(\overline{\psi},\lambda)\left(\frac{\sigma_\iota^2}{\sigma_\zeta^2+\sigma_\iota^2+e^{\overline{\psi}}+g_\sigma(\overline{\psi},\lambda)}\right)g_\mu(\overline{\psi},\lambda) \tag{2.39}$$

where $\kappa_{ya_j}(\overline{\psi},\lambda)$ denotes the use of private information at the A-SS.

In the next step, we log-linearize (2.16) around the A-SS:

$$\left(\frac{1+\epsilon}{1-\alpha}-1+\frac{1}{\theta}\right)\widehat{y}_{j,t} = \left(\frac{1+\epsilon}{1-\alpha}\right)a_{j,t}+\left(\frac{1}{\theta}-\gamma\right)\left(\int_{\mathcal{R}} E^{\omega_t}_{j,t,1}[\widehat{y}_t]\,\widetilde{f}_{j,t,1}(\omega_t)\,d\omega_t - H_y(\overline{\psi},\lambda)\right) \tag{2.40}$$

Therefore, matching coefficients lead to the following two equilibrium conditions

$$\left[\left(\frac{1+\epsilon}{1-\alpha}\right)-(1-\gamma)\right]\kappa_{ya_j}(\psi_t,\lambda) = \left(\frac{1+\epsilon}{1-\alpha}\right)-\left(\frac{1}{\theta}-\gamma\right)\left(\frac{\sigma_\iota^2}{\sigma_\zeta^2+\sigma_\iota^2+e^{\psi_t}+g_\sigma(\psi_t,\lambda)}\right)\kappa_{ya_j}(\psi_t,\lambda) \tag{2.41}$$

and

$$\left[\left(\frac{1+\epsilon}{1-\alpha}\right)-(1-\gamma)\right]\widehat{h}_y(\psi_t,\lambda) = \left(\frac{1}{\theta}-\gamma\right)H_y(\psi_t,\lambda)-\left(\frac{1}{\theta}-\gamma\right)H_y(\overline{\psi},\lambda) \tag{2.42}$$

by using the fact that under the proposed policy rules, we have that

$$\begin{aligned}\int_{\mathcal{R}} E^{\omega_t}_{j,t,1}[\widehat{y}_t]\,\widetilde{f}_{j,t,1}(\omega_t)\,d\omega_t = &\left[1-\left(\frac{\sigma_\iota^2}{\sigma_\zeta^2+\sigma_\iota^2+e^{\psi_t}+g_\sigma(\psi_t,\lambda)}\right)\right]\kappa_{ya_j}(\psi_t,\lambda)\,a_{j,t}\\ &+H_y(\widehat{\psi}_t,\lambda)-H_y(\overline{\psi},\lambda)\end{aligned}$$

where

$$H_y(\psi_t,\lambda) = \kappa_{ya_j}(\psi_t,\lambda)\left(\frac{\sigma_\iota^2}{\sigma_\zeta^2+\sigma_\iota^2+e^{\psi_t}+g_\sigma(\psi_t,\lambda)}\right)g_\mu(\psi_t,\lambda) \tag{2.43}$$

In what follows, we first present an auxiliary lemma that will be intensively used in later on proofs. It helps us prove that there exists a unique A-SS characterized by the use of private information $\overline{\kappa}_{ya_j} \equiv \kappa_{ya_j}(\overline{\psi},\lambda)$ and A-SS impacts of ambiguity shocks $\overline{h}_y(\overline{\psi})$. After proving the existence and uniqueness. of A-SS, we move on to prove that for any given amount of ambiguity ψ_t, there exists a unique $\kappa_{ya_j}(\psi_t,\lambda)$ manifesting the use of private information. And finally, existence and uniqueness of $\widehat{h}_y$ would be straight-forward given all the results we have established, which completes the whole proof.

Lemma 2.4. For any realized amount of ambiguity ψ_t, both $\kappa_{J\omega}(\psi_t,\lambda)$ and $\kappa_{J\omega\omega}(\psi_t,\lambda)$ are positive.

Proof. First of all, at the D-SS, it is straight-forward to show that $\chi(N^*)^{1+\epsilon} = (1-\alpha)(Y^*)^{1-\gamma}$. Then at the A-SS, it has to be the case that

$$\frac{(Y^*)^{1-\gamma}\left(1+(1-\gamma)\overline{h}_y\right)}{1-\gamma} - \chi\frac{(N^*)^{1+\epsilon}\left(1+(1+\epsilon)\overline{h}_n\right)}{1+\epsilon} > 0$$

Otherwise, it is better-off to be inactive by choosing $\overline{Y}=\overline{C}=\overline{N}=0$. Therefore, it must be the case that

$$\frac{\left(1+(1-\gamma)\overline{h}_y\right)}{1-\gamma} - \frac{(1-\alpha)\left(1+(1+\epsilon)\overline{h}_n\right)}{1+\epsilon} > 0 \tag{2.44}$$

Directly following (2.41), we know that $\kappa_{ya_j} > 0$. This implies that

$$(1+\epsilon)\kappa_{na_j} - (1-\gamma)\kappa_{ya_j} < 0 \tag{2.45}$$

since we know

$$(1+\epsilon)\kappa_{na_j} - (1-\gamma)\kappa_{ya_j} = -\left(\frac{1}{\theta}-\gamma\right)\left(\frac{\sigma_\iota^2}{\sigma_\zeta^2+\sigma_\iota^2+e^{\psi_t}+g_\sigma(\psi_t,\lambda)}\right)\kappa_{ya_j}$$

Furthermore, it can be shown that, by using $\chi(N^*)^{1+\epsilon} = (1-\alpha)(Y^*)^{1-\gamma}$, we would have that

$$\kappa_{J\omega}\left(\psi_t,\lambda\right) = \left(Y^*\right)^{1-\gamma}\left[\left(1+(1-\gamma)\overline{h}_y\right)\kappa_{ya_j}\left(\psi_t,\lambda\right) - \left(1+(1+\epsilon)\overline{h}_n\right)(1-\alpha)\kappa_{na_j}\left(\psi_t,\lambda\right)\right]$$

Then it is straight-forward to prove $\kappa_{J\omega}\left(\psi_t,\lambda\right) > 0$ given (2.44) and (2.45). Then we know $g_\mu\left(\psi_t,\lambda\right) < 0$, hence $H_y\left(\overline{\psi},\lambda\right) < 0$. Following (2.38), we will have that

$$(1-\gamma)\overline{h}_y > (1+\epsilon)\overline{h}_n \tag{2.46}$$

Since we know that

$$\kappa_{J\omega\omega}\left(\psi_t,\lambda\right) = \left(Y^*\right)^{1-\gamma}\left[\left(1+(1-\gamma)\overline{h}_y\right)(1-\gamma)\kappa^2_{ya_jt}\left(\psi_t,\lambda\right) - \left(1+(1+\epsilon)\overline{h}_n\right)(1-\alpha)(1+\epsilon)\kappa^2_{na_j}\left(\psi_t,\lambda\right)\right]$$

it is straight-forward to have $\kappa_{J\omega\omega}\left(\psi_t,\lambda\right) > 0$ given (2.45) and (2.46). ∎

Unique A-SS:

Define $S = \left(Y^*\right)^{1-\gamma}$, $X = \left[\frac{1+\epsilon}{1-\alpha} - (1-\gamma)\right]$, $\Sigma^2 = \frac{e^{\overline{\psi}}}{1+\lambda\overline{\kappa}_{J\omega\omega}}$ and finally $\beta_\iota = \frac{\sigma_\iota^2}{\sigma_\zeta^2+\sigma_\iota^2+\Sigma^2}$. Then A-SS can be characterized by the following four equations:

$$X\overline{\kappa}_{ya_j} = \left(\frac{1+\epsilon}{1-\alpha}\right) - \left(\frac{1}{\theta}-\gamma\right)\overline{\kappa}_{ya_j}\beta_\iota \tag{2.47}$$

$$X\overline{h}_y = -\left(\frac{1}{\theta}-\gamma\right)\overline{\kappa}_{ya_j}\beta_\iota\lambda\overline{\kappa}_{J\omega}\Sigma^2 \tag{2.48}$$

$$\overline{\kappa}_{J\omega} = S\left[\left(1+(1-\gamma)\overline{h}_y\right)\overline{\kappa}_{ya_j} - \left(1+(1+\epsilon)\overline{h}_n\right)(1-\alpha)\overline{\kappa}_{na_j}\right] \tag{2.49}$$

and

$$\overline{\kappa}_{J\omega\omega} = S\left[\left(1+(1-\gamma)\overline{h}_y\right)(1-\gamma)\overline{\kappa}^2_{ya_j} - \left(1+(1+\epsilon)\overline{h}_n\right)(1+\epsilon)(1-\alpha)\overline{\kappa}^2_{na_j}\right] \tag{2.50}$$

Focus on (2.47) first. Simple algebra leads to

$$\overline{\kappa}_{ya_j} = \frac{\frac{1+\epsilon}{1-\alpha}}{X+\left(\frac{1}{\theta}-\gamma\right)\beta_\iota} \equiv f\left(\overline{\kappa}_{ya_j},\overline{h}_y\right) \tag{2.51}$$

It can be shown that

$$\frac{\partial f}{\partial \overline{\kappa}_{ya_j}} < 0 \qquad\qquad \frac{\partial f}{\partial \overline{h}_y} < 0$$

For any given $\overline{h}_y$, the LHS of (2.47) is increasing in $\overline{\kappa}_{ya_j}$ and the RHS is decreasing in $\overline{\kappa}_{ya_j}$. Together with the boundary conditions, by intermediate value theorem, we can prove that for any $\overline{h}_y$ there exists a unique $\overline{\kappa}^*_{ya_j}\left(\overline{h}_y\right)$ that satisfy (2.51). Furthermore, we can show that such a function is decreasing in the sense that

$$\frac{d\overline{\kappa}_{ya_j}}{d\overline{h}_y} = \frac{\partial f/\partial \overline{h}_y}{1 - \partial f/\partial \overline{\kappa}_{ya_j}} < 0$$

Next focus on (2.48). Things are much more complicated here. Our target here is to show that given any identified function $\overline{\kappa}^*_{ya_j}\left(\overline{h}_y\right)$, there exists a unique $\overline{h}_y$ that satisfies (2.48). Existence can be easily proved by boundary conditions. Uniqueness can be ensured under some regularity conditions. Define LHS and RHS of (2.48) upon taking as $\overline{\kappa}^*_{ya_j}\left(\overline{h}_y\right)$ into consideration as $LHS(\overline{h}_y)$ and $RHS\left(\overline{h}_y\right)$. A sufficient condition for uniqueness is that

$$\frac{dRHS}{d\overline{h}_y} < \frac{dLHS}{d\overline{h}_y} = X \quad \text{whenever} \quad LHS = RHS$$

It can be shown that

$$\frac{dRHS}{d\overline{h}_y} = \lambda\overline{\kappa}_{J\omega}\Sigma^2\left[-\left(\frac{1}{\theta} - \gamma\right)\right]\frac{d\overline{\kappa}_{ya_j}\beta_\iota}{d\overline{h}_y} + \left[-\left(\frac{1}{\theta} - \gamma\right)\overline{\kappa}_{ya_j}\beta_\iota\right]\frac{d\lambda\overline{\kappa}_{J\omega}\Sigma^2}{d\overline{h}_y}$$

By using the fact that $\overline{\kappa}^*_{ya_j}\left(\overline{h}_y\right)$ is a solution for (2.47), we will have that

$$\left[-\left(\frac{1}{\theta} - \gamma\right)\right]\frac{d\overline{\kappa}_{ya_j}\beta_\iota}{d\overline{h}_y} = X\frac{d\overline{\kappa}_{ya_j}}{d\overline{h}_y}$$

Further evaluating it at $LHS = RHS$ implies that

$$\left[-\left(\frac{1}{\theta} - \gamma\right)\overline{\kappa}_{ya_j}\beta_\iota\right] = \frac{X\overline{h}_y}{\lambda\overline{\kappa}_{J\omega}\Sigma^2}$$

Therefore, we have that

$$\begin{aligned}\frac{1}{X}\frac{dRHS}{d\overline{h}_y} =&\lambda\overline{\kappa}_{J\omega}\Sigma^2\frac{d\overline{\kappa}_{ya_j}}{d\overline{h}_y}+\overline{h}_y\frac{d\log\left(\overline{\kappa}_{J\omega}\right)}{d\overline{h}_y}+\overline{h}_y\frac{d\log\left(\Sigma^2\right)}{d\overline{h}_y}\\ =&-\frac{X\overline{h}_y}{\left(\frac{1}{\theta}-\gamma\right)\overline{\kappa}_{ya_j}\beta_\iota}\frac{d\overline{\kappa}_{ya_j}}{d\overline{h}_y}+\overline{h}_y\frac{d\log\left(\overline{\kappa}_{J\omega}\right)}{d\overline{h}_y}+\overline{h}_y\frac{d\log\left(\Sigma^2\right)}{d\overline{h}_y}\end{aligned}$$

or equivalently

$$-\frac{1}{X\overline{h}_y}\frac{dRHS}{d\overline{h}_y} =\frac{X}{\left(\frac{1}{\theta}-\gamma\right)\overline{\kappa}_{ya_j}\beta_\iota}\frac{d\overline{\kappa}_{ya_j}}{d\overline{h}_y}-\frac{1}{\overline{\kappa}_{J\omega}}\frac{\partial\overline{\kappa}_{J\omega}}{\partial\overline{\kappa}_{ya_j}}\frac{d\overline{\kappa}_{ya_j}}{d\overline{h}_y}-\frac{1}{\overline{\kappa}_{J\omega}}\frac{\partial\overline{\kappa}_{J\omega}}{\partial\overline{h}_y}+\lambda\Sigma^2\frac{\partial\overline{\kappa}_{J\omega\omega}}{\partial\overline{h}_y}+\lambda\Sigma^2\frac{\partial\overline{\kappa}_{J\omega\omega}}{\partial\overline{\kappa}_{ya_j}}\frac{d\overline{\kappa}_{ya_j}}{d\overline{h}_y}$$

The followings can be easily proved:

$$\begin{aligned}\frac{\partial\overline{\kappa}_{J\omega}}{\partial\overline{\kappa}_{ya_j}} =&S\left[\left(1-\gamma\right)\overline{h}_y-\left(1+\epsilon\right)\overline{h}_n\right]=S\left(\frac{1}{\theta}-\gamma\right)\overline{\kappa}_{ya_j}\beta_\iota\lambda\overline{\kappa}_{J\omega}\Sigma^2\\ \frac{\partial\overline{\kappa}_{J\omega}}{\partial\overline{h}_y} =&S\left[\left(1-\gamma\right)\overline{\kappa}_{ya_j}-\left(1+\epsilon\right)\overline{\kappa}_{na_j}\right]=S\left(\frac{1}{\theta}-\gamma\right)\overline{\kappa}_{ya_j}\beta_\iota\end{aligned}$$

and

$$\begin{aligned}\frac{\partial\overline{\kappa}_{J\omega\omega}}{\partial\overline{\kappa}_{ya_j}} =&2S\left[\left(1+\left(1-\gamma\right)\overline{h}_y\right)\left(1-\gamma\right)\overline{\kappa}_{ya_j}-\left(1+\left(1+\epsilon\right)\overline{h}_n\right)\left(1+\epsilon\right)\overline{\kappa}_{na_j}\right]>0\\ \frac{\partial\overline{\kappa}_{J\omega\omega}}{\partial\overline{h}_y} =&S\left[\left(1-\gamma\right)^2\overline{\kappa}_{ya_j}^2-\left(1+\epsilon\right)^2\overline{\kappa}_{na_j}^2\right]>0\end{aligned}$$

when evaluating at A-SS, i.e., (2.47) and (2.48) are satisfied. Simple algebra would eventually imply

$$-\frac{1}{X\overline{h}_y}\frac{dRHS}{d\overline{h}_y} =A\frac{d\overline{\kappa}_{ya_j}}{d\overline{h}_y}+B$$

where we define the A and B are such that

$$\begin{aligned}A\equiv&\frac{X}{\left(\frac{1}{\theta}-\gamma\right)\overline{\kappa}_{ya_j}\beta_\iota}-S\left(\frac{1}{\theta}-\gamma\right)\overline{\kappa}_{ya_j}\beta_\iota\lambda\Sigma^2\\ &+2\lambda\Sigma^2S\left[\left(1+\left(1-\gamma\right)\overline{h}_y\right)\left(1-\gamma\right)\overline{\kappa}_{ya_j}-\left(1+\left(1+\epsilon\right)\overline{h}_n\right)\left(1+\epsilon\right)\overline{\kappa}_{na_j}\right]\\ =&\frac{X}{\left(\frac{1}{\theta}-\gamma\right)\overline{\kappa}_{ya_j}\beta_\iota}\\ &+\lambda\Sigma^2S\left[2\left(1+\left(1-\gamma\right)\overline{h}_y\right)\left(1-\gamma\right)\overline{\kappa}_{ya_j}-2\left(1+\left(1+\epsilon\right)\overline{h}_n\right)\left(1+\epsilon\right)\overline{\kappa}_{na_j}-\left(\frac{1}{\theta}-\gamma\right)\overline{\kappa}_{ya_j}\beta_\iota\right]\\ =&\frac{X}{\left(\frac{1}{\theta}-\gamma\right)\overline{\kappa}_{ya_j}\beta_\iota}+\lambda\Sigma^2S\left[2\left(1-\gamma\right)^2\overline{h}_y\overline{\kappa}_{ya_j}-2\left(1+\epsilon\right)\overline{h}_n\overline{\kappa}_{na_j}+\left(1-\gamma\right)\overline{\kappa}_{ya_j}-\left(1+\epsilon\right)\overline{\kappa}_{na_j}\right]>0\end{aligned}$$

and

$$
\begin{aligned}
B \equiv & -\frac{1}{\overline{\kappa}_{J\omega}}\frac{\partial\overline{\kappa}_{J\omega}}{\partial\overline{h}_y}+\lambda\Sigma^2\frac{\partial\overline{\kappa}_{J\omega\omega}}{\partial\overline{h}_y} \\
= & -\frac{1}{\overline{\kappa}_{J\omega}}\frac{\partial\overline{\kappa}_{J\omega}}{\partial\overline{h}_y}+\lambda\Sigma^2 S\left[(1-\gamma)^2\overline{\kappa}^2_{ya_j}-(1+\epsilon)^2\overline{\kappa}^2_{na_j}\right] \\
< & \frac{\lambda e^{\overline{\psi}}S\left[(1-\gamma)^2\overline{\kappa}^2_{ya_j}-(1+\epsilon)^2\overline{\kappa}^2_{na_j}\right]}{1+\lambda e^{\overline{\psi}}\overline{\kappa}_{J\omega\omega}}
\end{aligned}
$$

Therefore, we have that

$$
-\overline{h}_y B<\frac{\lambda e^{\overline{\psi}}S\left[-(1-\gamma)^2\overline{\kappa}^2_{ya_j}\overline{h}_y+(1+\epsilon)^2\overline{\kappa}^2_{na_j}(1-\alpha)\overline{h}_n\right]}{1+\lambda e^{\overline{\psi}}\overline{\kappa}_{J\omega\omega}}
$$

Finally, we can prove

$$
\begin{aligned}
& 1+\lambda e^{\overline{\psi}}\overline{\kappa}_{J\omega\omega}-\lambda e^{\overline{\psi}}S\left[-(1-\gamma)^2\overline{\kappa}^2_{ya_j}\overline{h}_y+(1+\epsilon)^2\overline{\kappa}^2_{na_j}(1-\alpha)\overline{h}_n\right] \\
= & 1+\lambda e^{\overline{\psi}}S\left[(1-\gamma)\overline{\kappa}^2_{ya_j}-(1+\epsilon)(1-\alpha)\overline{\kappa}^2_{na_j}+2(1-\gamma)^2\overline{\kappa}^2_{ya_j}\overline{h}_y-2(1+\epsilon)^2\overline{\kappa}^2_{na_j}(1-\alpha)\overline{h}_n\right] \\
= & 1+\lambda e^{\overline{\psi}}S\left\{\left[1+2(1-\gamma)\overline{h}_y\right](1-\gamma)\overline{\kappa}^2_{ya_j}-\left[1+2(1+\epsilon)\overline{h}_n\right](1+\epsilon)(1-\alpha)\overline{\kappa}^2_{na_j}\right\}>0
\end{aligned}
$$

where the last step can be justified by the fact that $\overline{h}_y$ is relatively small for any reasonable macro-application in the sense that $1+2(1-\gamma)\overline{h}_y>0$. Therefore, we arrive at the fact that

$$
-\overline{h}_y B<1
$$

which ensures uniqueness since it directly implies that $\frac{dRHS}{d\overline{h}_y}<X$.

Unique $\kappa_{ya_j}(\psi_t,\lambda)$:

Use of private information κ_{ya_j} is determined by (2.41). Denote the gap between LHS and RHS of (2.41) as $f\left(\kappa_{ya_j}\right)$ such that

$$
f\left(\kappa_{ya_j}\right)\equiv\left[\left(\frac{1+\epsilon}{1-\alpha}\right)-(1-\gamma)\right]\kappa_{ya_j}-\left(\frac{1+\epsilon}{1-\alpha}\right)+\left(\frac{1}{\theta}-\gamma\right)\left(\frac{\sigma^2_\iota}{\sigma^2_\zeta+\sigma^2_\iota+e^{\psi_t}+g_\sigma(\psi_t,\lambda)}\right)\kappa_{ya_j} \tag{2.52}
$$

It can be shown that

- $f\left(\kappa_{ya_j}\right)<0$ if $\kappa_{ya_j}<1$

- $f\left(\kappa_{ya_j}\right) > 0$ if $\kappa_{ya_j} > 1 + \frac{(1-\gamma)(1-\alpha)}{(1+\epsilon)-(1-\gamma)(1-\alpha)}$
- $f'\left(\kappa_{ya_j}\right) > 0$

The last item follows the fact that $g_\sigma(\psi_t, \lambda)$ is decreasing in κ_{ya_j} because $g_\sigma(\psi_t, \lambda)$ decreases in $\kappa_{J\omega\omega}$, which is an increasing function of κ_{ya_j}.

Finally, in the last step of the proof, it is straight-forward to demonstrate the existence and uniqueness for $\widehat{h}_y(\psi_t, \lambda)$ from (2.42) given existence and uniqueness given the existence and uniqueness for A-SS and κ_{ya_j}. ■

Proof of Proposition 2.2 . Directly follows the comparison between Proof of Proposition 2.1 and the solution for the beauty contest identified in the proposition. The comparative static analysis of g_μ and g_σ can be proved by demonstrating $\kappa_{J\omega}$ is increasing and $\kappa_{J\omega\omega}$. is decreasing in ψ_t.

■

Proof of Proposition 2.3 . It can be shown that (2.52) has the following properties regarding its partial derivatives evaluated at its equilibrium point $f\left(\kappa_{ya_j}\right) = 0$:

- $\frac{\partial f\left(\kappa_{ya_j}\right)}{\partial \kappa_{ya_j}}|_{f\left(\kappa_{ya_j}\right)=0} > 0$
- $\frac{\partial f\left(\kappa_{ya_j}\right)}{\partial \psi_t}|_{f\left(\kappa_{ya_j}\right)=0} < 0$

which are all straight-forward!. Then we can prove that $\kappa_{ya_j}(\psi_t, \lambda)$ is increasing in ψ_t. Then following the (2.27), $FD_t(\psi_t, \lambda)$ increases with ψ_t can be proved easily.

Also it can be shown that $\kappa_{J\omega}$ is increasing in ψ_t since it is increasing in κ_{ya_j}, which is an increasing function of ψ_t. Also note that we can transform (2.43) into

$$H_y(\psi_t, \lambda) = -\kappa_{ya_j}(\psi_t, \lambda)\left(\frac{\sigma_\iota^2}{\sigma_\zeta^2 + \sigma_\iota^2 + e^{\psi_t} + g_\sigma(\psi_t, \lambda)}\right)\left(e^{\psi_t} + g_\sigma(\psi_t, \lambda)\right)\lambda\kappa_{J\omega}$$

Following (2.41), we know that in equilibrium it must be the case that $e^{\psi_t} + g_\sigma(\psi_t, \lambda)$ is increasing in ψ_t. Then we know $H_y(\psi_t, \lambda)$ must be decreasing in ψ_t. Put this into (2.41), we complete the proof.

■

Proof of Lemma 2.3 and Proposition 2.5. First of all, it can be shown that

$$\begin{aligned}\left(\overline{\kappa}_{J\omega\omega}+\frac{1}{\lambda e^{\overline{\psi}}}\right)\frac{\partial\log\overline{\kappa}_{J\omega}}{\partial\overline{\kappa}_{ya_j}} =&\frac{1}{\overline{\kappa}_{J\omega}}\frac{\partial\overline{\kappa}_{J\omega}}{\partial\overline{\kappa}_{ya_j}}\left(\overline{\kappa}_{J\omega\omega}+\frac{1}{\lambda e^{\overline{\psi}}}\right)\\ =&S\left(\frac{1}{\theta}-\gamma\right)\overline{\kappa}_{ya_j}\beta_\iota\\ =&S\left[(1-\gamma)\kappa_{ya_j}-(1+\epsilon)\kappa_{na_j}\right]\\ <&2S\left[\left(1+(1-\gamma)\overline{h}_y\right)(1-\gamma)\overline{\kappa}_{ya_j}-\left(1+(1+\epsilon)\overline{h}_n\right)(1+\epsilon)\overline{\kappa}_{na_j}\right]\\ =&\frac{\partial\overline{\kappa}_{J\omega\omega}}{\partial\kappa_{ya_j}}\end{aligned}$$

The inequality follows the same argument as in the proof for unique A-SS. Then we can claim that

$$\frac{\partial\log\overline{\kappa}_{J\omega}}{\partial\overline{\kappa}_{ya_j}}<\frac{\lambda e^{\overline{\psi}}}{1+\lambda\overline{\kappa}_{J\omega\omega}e^{\overline{\psi}}}\frac{\partial\overline{\kappa}_{J\omega\omega}}{\partial\overline{\kappa}_{ya_j}}$$

At A-SS, we have that

$$g_\mu\left(\overline{\psi},\lambda\right)=-\frac{\lambda\overline{\kappa}_{J\omega}e^{\overline{\psi}}}{1+\lambda\overline{\kappa}_{J\omega\omega}e^{\overline{\psi}}}\Rightarrow\log\left(-g_\mu\left(\overline{\psi},\lambda\right)\right)=\log\left(\lambda\overline{\kappa}_{J\omega}e^{\overline{\psi}}\right)-\log\left(1+\lambda\overline{\kappa}_{J\omega\omega}e^{\overline{\psi}}\right)$$

Taking derivative leads to

$$\frac{\partial\log\left(-g_\mu\left(\overline{\psi},\lambda\right)\right)}{\partial\overline{\kappa}_{ya_j}}=\frac{\partial\log\overline{\kappa}_{J\omega}}{\partial\overline{\kappa}_{ya_j}}-\frac{\lambda e^{\overline{\psi}}}{1+\lambda\overline{\kappa}_{J\omega\omega}e^{\overline{\psi}}}\frac{\partial\overline{\kappa}_{J\omega\omega}}{\partial\overline{\kappa}_{ya_j}}<0$$

Therefore, we have that

$$\frac{\partial g_\mu\left(\overline{\psi},\lambda\right)}{\partial\overline{\kappa}_{ya_j}}>0 \tag{2.53}$$

At the A-SS when the amount of ambiguity is $\overline{\psi}$, equilibrium is characterized by (2.47) and (2.48). We can transform (2.48) by

$$X\overline{h}_y=\left[\left(\frac{1+\epsilon}{1-\alpha}\right)-X\overline{\kappa}_{ya_j}\right]g_\mu\left(\overline{\psi},\lambda\right) \tag{2.54}$$

As in the proof of unique A-SS, (2.47) defines a decreasing function of $\kappa_{yx}\left(\overline{h}_y\right)$. And an increase in σ_ι^2 or a decrease in σ_ζ^2 increases the RHS of (2.47) for any given $\overline{h}_y$. This indicates that the function $\kappa^*_{yx}\left(\overline{h}_y\right)$ derived from (2.47) shifted to the left in response to an increase in σ_ι^2 or a decrease in σ_ζ^2. Furthermore, the RHS of (2.54) increases with $\overline{\kappa}_{ya_j}$ marginally at the point where LHS equals RHS. Following the proof of unique A-SS, the gap between RHS and LHS of (2.54) is marginally increasing in $\overline{h}_y$ at the point where LHS equals RHS. Then a marginal increase in $\overline{\kappa}_{ya_j}$ implies a marginal increase in $\overline{y}$. Therefore, (2.54) defines an increasing function of $\kappa^{**}_{yx}\left(\overline{h}_y\right)$. The decreasing function $\kappa^{*}_{yx}\left(\overline{h}_y\right)$, the increasing function of $\kappa^{**}_{yx}\left(\overline{h}_y\right)$ and the left hand side shift of $\kappa^{*}_{yx}\left(\overline{h}_y\right)$ in combine predict that

$$\frac{d\overline{\kappa}_{ya_j}}{d\sigma^2_\iota} < 0 \qquad\qquad \frac{d\overline{\kappa}_{ya_j}}{d\sigma^2_{\sigma^2_\zeta}} > 0 \tag{2.55}$$

and

$$\frac{d\overline{h}_y}{d\sigma^2_\iota} < 0 \qquad\qquad \frac{d\overline{h}_y}{d\sigma^2_{\sigma^2_\zeta}} > 0 \tag{2.56}$$

Finally, (2.53) and (2.55) together prove Lemma 2.3. ■

2.B. Accuracy of the Approximation

Suppose that the allocation constitutes a Log-Normal equilibrium in the sense that in equilibrium $y_{j,t} \equiv \ln\left(Y_{j,t}\right)$ takes the following forms

$$y_{j,t} = \overline{y} + \kappa_{ya_j}\left(\psi_t, \lambda\right) a_{j,t} + \widehat{h}_y\left(\psi_t, \lambda\right)$$

Then we can, implement a quadratic approximation around the ambiguous steady state (A-SS) over $\overline{J}_t\left(\omega_t\right) \equiv E^{\omega_t}_{j,t,0}\left[\frac{Y_t^{1-\gamma}-1}{1-\gamma} - \chi\int_J \frac{\left(Y_{j,t}/A_{j,t}\right)^{(1+\epsilon)/(1-\alpha)}}{1+\epsilon} dj\right]$, which gives out the following quadratic functional forms with respect to ω_t:

$$\overline{J}_t\left(\omega_t\right) \approx const_t + \kappa_{J\omega}\left(\psi_t, \lambda\right)\omega_t + \frac{1}{2}\kappa_{J\omega\omega}\left(\psi_t, \lambda\right)\omega_t^2$$

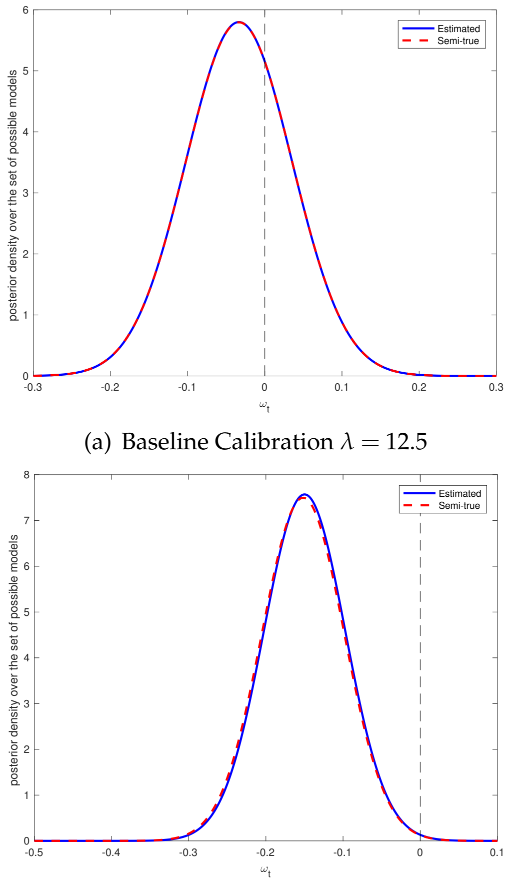
Directly following (2.17), the approximation implies a Normally distributed distorted posterior belief over the set of possible models. This is exactly what Lemma 2.1 is about. In addition, the fixed point condition (2.16) combined with normality in distorted belief over the set of possible models would automatically imply a conditional Log-Normal equilibrium such that

$$y_t = \overline{y} + \kappa_{ya_j}\left(\psi_t, \lambda\right)\int_J a_{j,t} dj + \widehat{h}_y\left(\psi_t, \lambda\right) + \underbrace{R_t + A_t + D_t}_{\text{uncertainty adjustment}}$$

where uncertainty adjustment consists of (a) risk adjustment R_t, (b) ambiguity adjustment and finally (c) dispersion adjustment. All of these adjustments are of second order impacts at the aggregates and have no impacts at all on the cyclical behaviours of belief divergence. Therefore, we ignore these adjustments in the approximated conditional Log-Normal equilibrium without loss of generality. Equivalently, this corresponds to the usual log-linearation of (2.16), but around the ambiguous steady state (A-SS) instead of deterministic steady state (D-SS). In sum, to approximate the conditional Log-Normal equilibrium, we implement a second order approximation over ex-ante expected value function of the representative household and at the same time a log-linearization of optimality conditions, both around the A-SS.

How accurate is our approximation? We address this issue into two steps. First of all, we highlight at what points of the analysis approximations are used. Then we analyze in details what are the nature of the approximations and conduct some evaluations.

During the characterization of the approximated conditional Log-Normal equilibrium, we first conduct a quadratic approximation of the ex-ante expected utility $\overline{J}_t(\omega_t)$ under a particular model ω_t so as to express the belief distortion in (2.17) into exponential-quadratic form. This implies approximated Normality for $\widetilde{f}_{j,t,1}(\omega_t)$ the distorted posterior over the set of possible models when evaluating marginal effects of an act. Second, given the approximated normality posterior, we log-linearize the optimality conditions, consisting of (2.1) and (2.15), around the A-SS. By doing so, we ignore uncertainty adjustments who are by nature a couple of second-order terms having negligible impacts on allocations and no impacts at all on belief divergence. For the second approximation, it is fairly standard in the literature. As a matter of fact, the ignored uncertainty adjustments are of magnitude $\sigma_\zeta^2 + \sigma_\iota^2 + e^{\psi_t}$. However, the first moment impacts of ambiguity shocks are of magnitude λe^{ψ_t}, which dominates uncertainty adjustments given the calibrated degree of ambiguity aversion λ. While, for the first approximation, we deal with this issue by comparing the estimated posterior density as in (2.17) and a semi-true posterior density numerically. To compute the semi-true posterior density, we take the estimated policy functions (2.18) and (2.19) as given and compute (2.17) analytically without using quadratic approximation over $\overline{J}_t(\omega_t)$. Figure 2.B.1 demonstrates such comparison where parameters are chosen such that $\theta = 1$, $\epsilon = 0.5$, $\alpha = 0.36$, $\overline{\psi} = -5.05$, $\sigma_\zeta = 0.0078$, $\sigma_\iota = 0.090$ and $\sigma_\tau = 0.47$. These are the same as our quantitative evaluations in Section 2.6. χ is chosen to be 1.895 to have hours in D-SS being 1/3, which



(a) Baseline Calibration $\lambda = 12.5$

(b) Extreme Calibration $\lambda = 100$

Figure 2.B.1. *Accuracy of the Approximation*

is the same standard when we calibrate χ in Section 2.6. Following Angeletos and La'O (2009), γ is chosen to be 0.2 to ensure an empirically plausible income effect of labor supply. Comparisons over estimated- and semi-true- posteriors are conducted under two parameterizations for degree of ambiguity aversion λ: one baseline calibration $\lambda = 12.5$ (the left-panel) and one counterfactually extreme calibration $\lambda = 100$ (the right-panel). Finally, we normalize the realization of island productivity by setting $a_{j,t} = 0$. This indicates the Bayesian posterior should be mean zero. Therefore, the leftwards shifts of all the four posterior densities manifest degree of pessimism due to ambiguity aversion. Finally, it turns out the approximation of the distorted posterior over the set of possible models is fairly accurate, not only for the baseline but also for the extreme calibration. Moreover, once we move to the extreme calibration, our approximation turns out to under-estimates both degree of pessimism and volatility in beliefs over the set of possible models when comparing to the semi-true counter-part.

This further indicates that our approximation is quite conservative under extreme calibrations.

2.C. Labor Wedges

Define the marginal rate of intra-temporal substitution between leisure and consumption $MRSN_{j,t}$ and labor productivity $MPL_{j,t}$ on island j by

$$MRSN_{j,t} = \epsilon \widehat{n}_{j,t} + \widehat{c}_t$$

and

$$MPL_{j,t} = \widehat{y}_{j,t} - \widehat{n}_{j,t}$$

Then we can define island j labor wedge from the perspective of household by the gap between real wage $\widehat{w}_{j,t}$ and island j marginal rate of intra-temporal substitution between leisure and consumption $MRSN_{j,t}$:

$$\tau^{nh}_{j,t} \equiv \widehat{w}_{j,t} - MRSN_{j,t}$$

and island j labor wedge from the perspective of firm as the gap between island j labor productivity $MPL_{j,t}$ and real wage $\widehat{w}_{j,t}$:

$$\tau^{nf}_{j,t} \equiv MPL_{j,t} - \widehat{w}_{j,t}$$

Finally, we define the total labor wedge in the economy by the cross-sectional average over the sum of the two:

$$\tau^{n}_{t} \equiv \int_J \left(\tau^{nh}_{j,t} + \tau^{nh}_{j,t} \right) dj$$

Plugging in optimality conditions for labor, model implied wedges are such that

$$\begin{aligned}
\tau^{nh}_{j,t} &= \int_{\mathcal{R}} E^{\omega_t}_{j,t,1} \left[\widehat{c}_t \right] \widetilde{f}_{j,t,1} \left(\omega_t \right) d\omega_t - \widehat{c}_t \\
\tau^{nf}_{j,t} &= \frac{1}{\theta} \left[\widehat{y}_{j,t} - \int_{\mathcal{R}} E^{\omega_t}_{j,t,1} \left[\widehat{y}_t \right] \widetilde{f}_{j,t,1} \left(\omega_t \right) d\omega_t \right]
\end{aligned}$$

2.D. Details for Quantitative Analysis

2.D.1. Equilibrium Conditions

Equilibrium can be characterized by the standard TVC and following conditions:

- labor optimality condition

$$\chi N_{j,t}^{\epsilon} = \left(\int_{\mathcal{R}} E_{j,t,1}^{\omega_t} \left[\frac{1}{C_t} \left(\frac{Y_{j,t}}{Y_t} \right)^{-\frac{1}{\theta}} \right] \widetilde{f}_{j,t,1} (\omega_t)\, d\omega_t \right) \left((1-\alpha) \frac{Y_{j,t}}{N_{j,t}} \right)$$

- optimal capital demand condition

$$R_{j,t} \int_{\mathcal{R}} E_{j,t,1}^{\omega_t} \left[\frac{1}{C_t} \right] \widetilde{f}_{j,t,1} (\omega_t)\, d\omega_t = \left(\int_{\mathcal{R}} E_{j,t,1}^{\omega_t} \left[\frac{1}{C_t} P_{j,t} \right] \widetilde{f}_{j,t,1} (\omega_t)\, d\omega_t \right) \left(\alpha \frac{Y_{j,t}}{K_{j,t}} \right)$$

- Euler equation

$$\frac{1}{C_t} = \beta \int_{\mathcal{R}} E_{t,2}^{\omega_{t+1}} \left[\frac{1}{C_{t+1}} \left((1-\delta) + R_{j,t+1} \right) \right] \widetilde{f}_{t,2} (\omega_{t+1})\, d\omega_{t+1} \tag{2.57}$$

- budget constraint

$$Y_t = C_t + \int_J I_{j,t} dj \tag{2.58}$$

- capital accumulation

$$K_{j,t+1} = (1-\delta) K_{j,t} + I_{j,t} \tag{2.59}$$

- production function for island commodities

$$Y_{j,t} = A_{j,t} K_{j,t}^{\alpha} N_{j,t}^{1-\alpha} \tag{2.60}$$

- production function for final goods

$$\log Y_t = \int_J \log Y_{j,t} dj \tag{2.61}$$

- value function recursion for $J_t \equiv J\left(\left\{K_{j,t}\right\}, a_{t-1}, z_t, \zeta_t, \psi_t\right)$:

$$J_t = \max_{C_t, \left\{K_{j,t+1}\right\}} \ln\left(C_t\right) + \beta\left(-\frac{1}{\lambda}\right)\ln\left(\int_{\mathcal{R}} e^{-\lambda E_{t,2}^{\omega_{t+1}}\left[J_{t+1}\right]} f_t\left(\omega_{t+1}\right) d\omega_{t+1}\right)$$

- distorted beliefs $\left\{\widetilde{f}_{j,t,1}\left(\omega_t\right), \widetilde{f}_{t,2}\left(\omega_{t+1}\right)\right\}$ at stage 1 and at stage 2 respectively

$$\widetilde{f}_{j,t,1}\left(\omega_t\right) = e^{-\lambda E_{t,0}^{\omega_t}\left[J_t\right]} f\left(x_{j,t} | \omega_t\right) f_t\left(\omega_t\right)$$

and

$$\widetilde{f}_{t,2}\left(\omega_{t+1}\right) = e^{-\lambda E_{t,2}^{\omega_t}\left[J_{t+1}\right]} f_t\left(\omega_{t+1}\right)$$

Observe that in equilibrium, it will be the case that $K_{j,t} = K_t$ for all $t > 0$. This is because at stage 2 there exists no heterogeneity in belief over capital return $r_{j,t}$ across islands. Therefore, capital supply exhibits no heterogeneity. In what follows, we replace all $K_{j,t}$ with K_t for simplicity.

2.D.2. Solution Method

We propose the following semi-linear policy rules of conditional log-normal equilibrium for island employment $\widehat{n}_{j,t}$, output $\widehat{y}_{j,t}$, wage rate $\widehat{w}_{j,t}$ and rental rental rate of capital $\widehat{r}_{j,t}$ at stage 1 of period t:

$$\begin{aligned}
\widehat{y}_{j,t} &= \kappa_{yk}\widehat{k}_t + \kappa_{ya}a_{t-1} + \kappa_{yx}\left(\psi_t\right)x_{j,t} + \widehat{h}_y\left(\psi_t\right)\\
\widehat{n}_{j,t} &= \kappa_{nk}\widehat{k}_t + \kappa_{na}a_{t-1} + \kappa_{nx}\left(\psi_t\right)x_{j,t} + \widehat{h}_n\left(\psi_t\right)\\
\widehat{w}_{j,t} &= \kappa_{wk}\widehat{k}_t + \kappa_{wa}a_{t-1} + \kappa_{wx}\left(\psi_t\right)x_{j,t} + \widehat{h}_w\left(\psi_t\right)\\
\widehat{r}_{j,t} &= \kappa_{rk}\widehat{k}_t + \kappa_{ra}a_{t-1} + \kappa_{rx}\left(\psi_t\right)x_{j,t} + \widehat{h}_r\left(\psi_t\right)
\end{aligned}$$

and the for consumption $\widehat{c}_t$, investment $\widehat{i}_t$ and capital stock tomorrow $\widehat{k}_{t+1}$ at stage 2 of period t:

$$\begin{aligned}
\widehat{c}_t &= \kappa_{ck}\widehat{k}_t + \kappa_{ca}a_{t-1} + \kappa_{cz}\left(\psi_t\right)z_t + \kappa_{c\zeta}\left(\psi_t\right)\zeta_t + \widehat{h}_c\left(\psi_t\right)\\
\widehat{i}_t &= \kappa_{ik}\widehat{k}_t + \kappa_{ia}a_{t-1} + \kappa_{iz}\left(\psi_t\right)z_t + \kappa_{i\zeta}\left(\psi_t\right)\zeta_t + \widehat{h}_i\left(\psi_t\right)\\
\widehat{k}_{t+1} &= \kappa_{kk}\widehat{k}_t + \kappa_{ka}a_{t-1} + \kappa_{kz}\left(\psi_t\right)z_t + \kappa_{k\zeta}\left(\psi_t\right)\zeta_t + \widehat{h}_k\left(\psi_t\right)
\end{aligned}$$

Quadratic approximation of value function

Under the proposed policy rules, quadratic approximation of household period utility around the A-SS is such that

$$\ln\left(C_t\right) - \chi \int_j \frac{N_{j,t}^{1+\epsilon}}{1+\epsilon} dj \tag{2.62}$$

$$\approx c^* + \overline{h}_c + \kappa_{ck}\widehat{k}_t + \kappa_{ca}a_{t-1} + \kappa_{cz,t}z_t + \kappa_{c\zeta,t}\zeta_t + \widehat{h}_c\left(\psi_t\right) \tag{2.63}$$

$$-\chi \frac{\overline{N}^{1+\epsilon}}{1+\epsilon} \int_j \Big[1 + (1+\epsilon)\left(\kappa_{nk}\widehat{k}_t + \kappa_{na}a_{t-1} + \kappa_{nx,t}x_{j,t} + \widehat{h}_n\left(\psi_t\right)\right) \tag{2.64}$$

$$+\frac{1}{2}(1+\epsilon)^2\left(\kappa_{nk}\widehat{k}_t + \kappa_{na}a_{t-1} + \kappa_{nx,t}x_{j,t} + \widehat{h}_n\left(\psi_t\right)\right)^2 \Big] dj \tag{2.65}$$

$$\tag{2.66}$$

Guess that J_t is given by

$$J_t = J^* + \overline{h}_J + \kappa_{Jk}\widehat{k}_t + \kappa_{Ja}a_{t-1} + \kappa_{Jz,t}z_t + \kappa_{J\zeta,t}\zeta_t \tag{2.67}$$

$$+ \kappa_{Jka}\widehat{k}_t a_{t-1} + \kappa_{Jkz,t}\widehat{k}_t z_t + \kappa_{Jk\zeta,t}\widehat{k}_t\zeta_t + \kappa_{Jaz,t}a_{t-1}z_t + \kappa_{Ja\zeta,t}a_{t-1}\zeta_t + \kappa_{Jz\zeta}z_t\zeta_t \tag{2.68}$$

$$+ \frac{1}{2}\kappa_{Jkk}\widehat{k}_t^2 + \frac{1}{2}\kappa_{Jaa}a_{t-1}^2 + \frac{1}{2}\kappa_{Jzz,t}z_t^2 + \frac{1}{2}\kappa_{J\zeta\zeta,t}\zeta_t^2 + \widehat{h}_J\left(\psi_t\right) \tag{2.69}$$

Hence we have

$$\left(-\frac{1}{\lambda}\right)\ln\left(\int_{\omega_{t+1}} e^{-\lambda E_{t,2}^{\omega_{t+1}}[J_{t+1}]} \widetilde{f}_{t,2}^{c}\left(\omega_{t+1}\right) d\omega_{t+1}\right) \tag{2.70}$$

$$= constant_t + \kappa_{Jk}\widehat{k}_{t+1} + \kappa_{Ja}a_t + \kappa_{Jka}\widehat{k}_{t+1}a_t + \frac{1}{2}\kappa_{Jkk}\widehat{k}_{t+1}^2 + \frac{1}{2}\kappa_{Jaa}a_t^2 \tag{2.71}$$

Value function recursion implies that

$$E_{t,0}^{\omega_t}\left[J_t\right] = \left(\widetilde{J^* + \overline{h}_J + \widehat{h}_J\left(\psi_t\right)}\right) + \kappa_{Jk}\widehat{k}_t + \kappa_{Ja}a_{t-1} + \kappa_{Jz,t}\omega_t + \kappa_{Jka}\widehat{k}_t a_{t-1} + \kappa_{Jkz,t}\widehat{k}_t\omega_t + \kappa_{Jaz,t}a_{t-1}\omega_t \tag{2.72}$$

$$+ \frac{1}{2}\kappa_{Jkk}\widehat{k}_t^2 + \frac{1}{2}\kappa_{Jaa}a_{t-1}^2 + \frac{1}{2}\kappa_{Jzz,t}\omega_t^2 \tag{2.73}$$

$$\approx constant_t + \kappa_{Jz,t}\omega_t + \frac{1}{2}\kappa_{Jzz,t}\omega_t^2 \tag{2.74}$$

where the following matching coefficients leads to

$$\kappa_{Jz,t} = \kappa_{cz,t} - \chi N^* \left[1 + (1+\epsilon)\overline{h}_n\right] \kappa_{nx,t} + \beta \kappa_{Jk} \kappa_{kz,t} \tag{2.75}$$

$$\kappa_{Jk} = \kappa_{ck} - \chi N^* \left[1 + (1+\epsilon)\overline{h}_n\right] \kappa_{nk} + \beta \kappa_{Jk} \kappa_{kk} \tag{2.76}$$

$$\kappa_{Jzz,t} = -(1+\epsilon)\chi N^* \left[1 + (1+\epsilon)\overline{h}_n\right] \kappa_{nx,t}^2 + \beta \kappa_{Jkk} \kappa_{kz,t}^2 \tag{2.77}$$

$$\kappa_{Jkk} = -(1+\epsilon)\chi N^* \left[1 + (1+\epsilon)\overline{h}_n\right] \kappa_{nk}^2 + \beta \kappa_{Jkk} \kappa_{kk}^2 \tag{2.78}$$

To arrive these expression, we first approximate $e^{(1+\epsilon)\overline{h}_n}$ by $1 + (1+\epsilon)\overline{h}_n$ and by ignoring a couple of higher order terms without loss of generality.

Furthermore, we can approximate by

$$E_{t,2}^{\omega_t}\left[J_{t+1}\right] \approx constant_t + \overline{\kappa_{Jz,t+1}}\omega_t + \frac{1}{2}\overline{\kappa_{Jzz,t+1}}\omega_t^2$$

where term with over-line denotes period t expectation over that term in period $t+1$.

Distorted posterior beliefs

Distorted beliefs are normal with the following kernels

$$\widetilde{f}_{j,t,1}(\omega_t) = e^{-\lambda\left(\kappa_{Jz,t}\omega_t + \frac{1}{2}\kappa_{Jzz,t}\omega_t^2\right)} f\left(x_{j,t}|\omega_t\right) f_t(\omega_t)$$

and

$$\widetilde{f}_{t,2}(\omega_{t+1}) = e^{-\lambda\left(\overline{\kappa_{Jz,t+1}}\omega_t + \frac{1}{2}\overline{\kappa_{Jzz,t+1}}\omega_t^2\right)} f_t(\omega_{t+1})$$

Ambiguous Steady State

A-SS can be characterized by

$$\begin{aligned}
\log(\chi) + (1+\epsilon)n^* + (1+\epsilon)\overline{h}_n =& \log(1-\alpha) + y^* - c^* + \overline{h}_y - \overline{h}_c + H_n\left(\overline{\psi}\right) \\
y^* + \overline{h}_y =& (1-\alpha)\left(n^* + \overline{h}_n\right) + \alpha\left(k^* + \overline{h}_k\right) \\
r^* + \overline{h}_r =& \log(\alpha) + y^* - k^* + \overline{h}_y - \overline{h}_k + H_r\left(\overline{\psi}\right) \\
1 =& O_1 + O_2 \\
X_c + X_i =& X_y \\
X_i =& \delta X_k
\end{aligned}$$

where terms with star denote the D-SS and the auxiliary functions are such that

$$
\begin{aligned}
H_n(\psi_t) =& \left[\left(\frac{1}{\theta}\kappa_{yx,t} - \kappa_{cz,t}\right)\left(\frac{\sigma_\iota^2}{\sigma_\zeta^2+\sigma_\iota^2}\right) + \kappa_{c\zeta,t}\left(\frac{\sigma_\zeta^2}{\sigma_\zeta^2+\sigma_\iota^2}\right)\right]\left(\frac{-\lambda\kappa_{Jz,t}}{\frac{1}{\sigma_\zeta^2+\sigma_\iota^2}+\frac{1}{e^{\psi_t}}+\lambda\kappa_{Jzz,t}}\right)\\
H_r(\psi_t) =& \frac{1}{\theta}\kappa_{yx,t}\left(\frac{\sigma_\iota^2}{\sigma_\zeta^2+\sigma_\iota^2}\right)\left(\frac{-\lambda\kappa_{Jz,t}}{\frac{1}{\sigma_\zeta^2+\sigma_\iota^2}+\frac{1}{e^{\psi_t}}+\lambda\kappa_{Jzz,t}}\right)\\
S_1(\psi_t) \equiv& \left(\overline{\kappa_{rx,t+1}-\kappa_{cz,t+1}}\right)\left(\frac{-\lambda\overline{\kappa_{Jz,t+1}}}{\frac{1}{e^{\psi_{t,t+1}}}+\lambda\overline{\kappa_{Jzz,t+1}}}\right) + \left(\overline{\widehat{h}_r(\psi_{t+1}) - \widehat{h}_c(\psi_{t+1})}\right)\\
S_2(\psi_t) \equiv& -\left(\overline{\kappa_{cz,t+1}}\right)\left(\frac{-\lambda\overline{\kappa_{Jz,t+1}}}{\frac{1}{e^{\psi_{t,t+1}}}+\lambda\overline{\kappa_{Jzz,t+1}}}\right) - \overline{\widehat{h}_c(\psi_{t+1})}\\
O_1 =& \beta e^{r^*+\overline{h}_r+S_1(\overline{\psi})}\\
O_2 =& \beta(1-\delta)e^{S_2(\overline{\psi})}
\end{aligned}
$$

and

$$
X_c = \left(e^{c^*+\overline{h}_c}\right) \qquad X_i = \left(e^{i^*+\overline{h}_i}\right) \qquad X_y = \left(e^{y^*+\overline{h}_y}\right) \qquad X_k = \left(e^{k^*+\overline{h}_k}\right)
$$

Log-linearization of optimality conditions

Log-linearzation of optimality conditions around the A-SS leads to

- labor optimality

$$
(1+\epsilon)\widehat{n}_{j,t} = \left(1-\frac{1}{\theta}\right)\widehat{y}_{j,t} + \int_{\omega_t} E_{j,t,1}^{\omega_t}\left[\frac{1}{\theta}\widehat{y}_t - \widehat{c}_t\right]\widetilde{f}_{j,t,1}(\omega_t)\,d\omega_t - H_n(\overline{\psi})
$$

- optimal capital demand condition

$$
\widehat{r}_{j,t} = \left(1-\frac{1}{\theta}\right)\widehat{y}_{j,t} + \int_{\omega_t} E_{j,t,1}^{\omega_t}\left[\frac{1}{\theta}\widehat{y}_t - \widehat{k}_t\right]\widetilde{f}_{j,t,1}(\omega_t)\,d\omega_t - H_r(\overline{\psi})
$$

- Euler equation

$$
c_t = O_1\int_{\mathcal{R}} E_{t,2}^{\omega_{t+1}}\left[\widehat{r}_{j,t+1} - \widehat{c}_{t+1}\right]\widetilde{f}_{t,2}(\omega_{t+1})\,d\omega_{t+1} - O_2\int_{\mathcal{R}} E_{t,2}^{\omega_{t+1}}\left[\widehat{c}_{t+1}\right]\widetilde{f}_{t,2}(\omega_{t+1})\,d\omega_{t+1}
$$

- budget constraint

$$
X_c\widehat{c}_t + X_i\widehat{i}_t = X_y\widehat{y}_t
$$

- capital accumulation

$$X_k\widehat{k}_{t+1} = (1-\delta)\, X_k\widehat{k}_t + X_i\widehat{i}_t$$

- production function for island commodities

$$\widehat{y}_{j,t} = \rho a_{t-1} + x_{j,t} + \alpha\widehat{k}_t + (1-\alpha)\,\widehat{n}_{j,t}$$

- production function for final goods

$$\widehat{y}_t = \int_J \widehat{y}_{j,t} dj$$

Systems of undetermined coefficients

Combining the proposed policy rules and log-linearized optimality conditions, we arrive at the following systems of undetermined coefficients:

- Determination of κ_{J*}

$$\begin{aligned}
\kappa_{Jz,t} =& \kappa_{cz,t} - \chi\overline{N}^{1+\epsilon}\kappa_{nx,t} + \beta\kappa_{Jk}\kappa_{kz,t} \\
\kappa_{Jk} =& \kappa_{ck} - \chi\overline{N}^{1+\epsilon}\kappa_{nk} + \beta\kappa_{Jk}\kappa_{kk} \\
\kappa_{Jzz,t} =& -(1+\epsilon)\,\chi\overline{N}^{1+\epsilon}\kappa_{nx,t}^2 + \beta\kappa_{Jkk}\kappa_{kz,t}^2 \\
\kappa_{Jkk} =& -(1+\epsilon)\,\chi\overline{N}^{1+\epsilon}\kappa_{nk}^2 + \beta\kappa_{Jkk}\kappa_{kk}^2
\end{aligned}$$

- Determination of ambiguous steady state

$$\begin{aligned}
\log(\chi) + (1+\epsilon)\, n^* + (1+\epsilon)\,\overline{h}_n =& \log(1-\alpha) + y^* - c^* + \overline{h}_y - \overline{h}_c + H_n\left(\overline{\psi}\right) \\
y^* + \overline{h}_y =& (1-\alpha)\left(n^* + \overline{h}_n\right) + \alpha\left(k^* + \overline{h}_k\right) \\
r^* + \overline{h}_r =& \log(\alpha) + y^* - k^* + \overline{h}_y - \overline{h}_k + H_r\left(\overline{\psi}\right) \\
1 =& O_1 + O_2 \\
X_c + X_i =& X_y \\
X_i =& \delta X_k
\end{aligned}$$

- Determination of κ_{*k} and κ_{*a}

$$
\begin{aligned}
(1+\epsilon)\,\kappa_{nk} =& \left(1-\frac{1}{\theta}\right)\kappa_{yk} + \left[\frac{1}{\theta}\kappa_{yk} - \kappa_{ck}\right] \\
(1+\epsilon)\,\kappa_{na} =& \left(1-\frac{1}{\theta}\right)\kappa_{ya} + \left[\frac{1}{\theta}\kappa_{ya} - \kappa_{ca}\right] \\
\kappa_{yk} =& (1-\alpha)\,\kappa_{nk} + \alpha \\
\kappa_{ya} =& \rho + (1-\alpha)\,\kappa_{na} \\
\kappa_{rk} =& \kappa_{yk} - 1 \\
\kappa_{ra} =& \kappa_{ya} \\
-\kappa_{ck} =& (O_1\kappa_{rk} - \kappa_{ck})\,\kappa_{kk} \\
-\kappa_{ca} =& (O_1\kappa_{rk} - \kappa_{ck})\,\kappa_{ka} + (O_1\kappa_{ra} - \kappa_{ca})\,\rho \\
X_c\kappa_{ck} + X_i\kappa_{ik} =& X_y\kappa_{yk} \\
X_c\kappa_{ca} + X_i\kappa_{ia} =& X_y\kappa_{ya} \\
\kappa_{kk} =& (1-\delta) + \delta\kappa_{ik} \\
\kappa_{ka} =& \delta\kappa_{ia}
\end{aligned}
$$

- Determination of $\kappa_{*x,t}$, $\kappa_{*z,t}$ and $\kappa_{*\omega,t}$

$$
\begin{aligned}
(1+\epsilon)\,\kappa_{nx,t} =& \left(1-\frac{1}{\theta}\right)\kappa_{yx,t} + \left[\left(\frac{1}{\theta}\kappa_{yx,t} - \kappa_{cz,t}\right)\left(\frac{\sigma_\zeta^2}{\sigma_\zeta^2+\sigma_\iota^2}\right) - \kappa_{c\zeta,t}\left(\frac{\sigma_\zeta^2}{\sigma_\zeta^2+\sigma_\iota^2}\right)\right] \\
&+ \left[\left(\frac{1}{\theta}\kappa_{yx,t} - \kappa_{cz,t}\right)\left(\frac{\sigma_\iota^2}{\sigma_\zeta^2+\sigma_\iota^2}\right) + \kappa_{c\zeta,t}\left(\frac{\sigma_\zeta^2}{\sigma_\zeta^2+\sigma_\iota^2}\right)\right]\left(\frac{\frac{1}{\sigma_\zeta^2+\sigma_\iota^2}}{\frac{1}{\sigma_\zeta^2+\sigma_\iota^2} + \frac{1}{e^{\psi_t}} + \lambda\kappa_{Jzz,t}}\right) \\
\kappa_{yx,t} =& 1 + (1-\alpha)\,\kappa_{nx,t} \\
\kappa_{rx,t} =& \left(1-\frac{1}{\theta}\right)\kappa_{yx,t} + \frac{1}{\theta}\kappa_{yx,t}\left(\frac{\sigma_\zeta^2}{\sigma_\zeta^2+\sigma_\iota^2}\right) \\
&+ \left(\frac{1}{\theta}\kappa_{yx,t}\right)\left(\frac{\sigma_\iota^2}{\sigma_\zeta^2+\sigma_\iota^2}\right)\left(\frac{\frac{1}{\sigma_\zeta^2+\sigma_\iota^2}}{\frac{1}{\sigma_\zeta^2+\sigma_\iota^2} + \frac{1}{e^{\psi_t}} + \lambda\kappa_{Jzz}}\right) \\
-\kappa_{cz,t} =& (O_1\kappa_{rk} - \kappa_{ck})\,\kappa_{kz,t} \\
-\kappa_{c\zeta,t} =& (O_1\kappa_{rk} - \kappa_{ck})\,\kappa_{k\zeta,t} + (O_1\kappa_{ra} - \kappa_{ca}) \\
X_c\kappa_{cz,t} + X_i\kappa_{iz,t} =& X_y\kappa_{yx,t} \\
X_c\kappa_{c\zeta,t} + X_i\kappa_{i\zeta,t} =& 0 \\
\kappa_{kz,t} =& \delta\kappa_{iz,t} \\
\kappa_{k\zeta,t} =& \delta\kappa_{i\zeta,t}
\end{aligned}
$$

- Determination of $\widehat{h}_*(\psi_t)$

$$
\begin{aligned}
(1+\epsilon)\widehat{h}_n(\psi_t) &= \left(1-\frac{1}{\theta}\right)\widehat{h}_y(\psi_t) + \left[\frac{1}{\theta}\widehat{h}_y(\psi_t) - \widehat{h}_c(\psi_t)\right] + \widehat{H}_n(\psi_t) \\
\widehat{h}_y(\psi_t) &= (1-\alpha)\widehat{h}_n(\psi_t) \\
\widehat{h}_r(\psi_t) &= \widehat{h}_y(\psi_t) + \widehat{H}_r(\psi_t) \\
-\widehat{h}_c(\psi_t) &= (O_1\kappa_{rk} - \kappa_{ck})\widehat{h}_k(\psi_t) + O_1\widehat{S}_1(\psi_t) + O_2\widehat{S}_2(\psi_t) \\
X_c\widehat{h}_c(\psi_t) + X_i\widehat{h}_i(\psi_t) &= X_y\widehat{h}_y(\psi_t) \\
\widehat{h}_k(\psi_t) &= \delta\widehat{h}_i(\psi_t)
\end{aligned}
$$

where $\widehat{H}_n(\psi_t) \equiv H_n(\psi_t) - H_n(\bar{\psi})$ and $\widehat{H}_r(\psi_t) \equiv H_r(\psi_t) - H_r(\bar{\psi})$.

The algorithm

To quantitatively solve for the equilibrium, we first discretize the ambiguity process by a stationary Markov process with 11 states. This transforms functions over ψ_t into 11 undetermined coefficients. Then we arrive at a non-linear large system of equations regarding undetermined coefficients. We describe the algorithm to solve this system of undermined coefficients below.

1. Guess that the ambiguous steady state coincides with deterministic steady state such that $\bar{h}_S = 0$ for all $S \in \{y,c,n,i,k,r\}$. This leads to $O_1 = \beta e^{r^*+\bar{h}_r}$ and $O_2 = \beta(1-\delta)$. Further guess that $S_1(\overline{\psi}) = S_1(\overline{\psi}) = 0$.
2. Given $\{O_1,O_2,S_1,S_2,X_c,X_y,X_i,X_k\}$, solve for κ_{*k} and κ_{*a}. We restricts $\kappa_{kk} < 1$ to ensure TVC being satisfied.
3. Jointly solve κ_{J*}, $\kappa_{*x,t}$, $\kappa_{*z,t}$ and $\kappa_{*\omega,t}$ given the solution for κ_{*k} and κ_{*a}.
4. Solve $\widehat{h}_*(\psi_t)$ for the given solution for κ_{J*}, $\kappa_{*x,t}$, $\kappa_{*z,t}$, $\kappa_{*\omega,t}$, κ_{*k} and κ_{*a}.
5. Solve for the A-SS and compute new levels of $\{O_1,O_2,S_1,S_2,X_c,X_y,X_i,X_k\}$.
6. Check for convergence over $\{O_1,O_2,S_1,S_2,X_c,X_y,X_i,X_k\}$. If converge, stop. Otherwise, go back to Step 2.

Convergence can be achieved when the amount of ambiguity at A-SS $\overline{\psi}$ is not extremely high or when degree of ambiguity aversion λ is not extremely high. This is because these are the conditions to ensure unique A-SS as in the simple model without capital. Other than this constraint, the algorithm works well in the approximation of the semi-linear conditional log-normal equilibrium.

Chapter 3

Financial Sentiments and Coordinated Information Provision[1]

[1] I thank Fabrizio Zilibotti, Heng Chen, Nir Jaimovich and Pengfei Wang for their useful discussions and comments.

3.1. Introduction

Empirical studies have documented that financial market fluctuations are disconnected with market fundamentals. For example, Shiller (2003) states that financial markets are inefficient because fundamentals are not sufficient to rationalize stock market volatility. Hall (2001) also finds that many high tech firms with sustained high stock market value turned out to have negative earnings. Further the non-fundamental volatilities are shown to reflect financial sentiment. For example, Verma and Verma (2007) empirically finds "significant positive effects of investor sentiments on stock returns."[2]

Where does the sentiment come from in financial market? What are the origins of financial sentiment? Financial markets are blamed for its inefficiency in providing the "right" amount of information over market fundamentals, which leads to waves of pessimism and optimism preventing the market from working efficiently. In this paper, we show that financial sentiments may stem from the process of information production and transmission. Specifically, the strategic interaction between information providers, e.g., analysts, and information receivers, e.g., investors, can be the driving force of sentiment or non-fundament fluctuations.

Framework. We formalize such an idea in a sender-receiver game that features endogenous information production and transmission. Analysts offer financial forecasts to investors. Specifically, analysts draft reports that analyze the strength of market fundamental and aggregate investment, e.g., earnings and multiples of the stock price. Their reports offer a forecast or recommendation based on their estimates of both fundamental and aggregate actions (eg. average recommendation made by other analysts and level of aggregate investment). Investors make investment decisions, based on the information they acquire from the analysts. Their payoff depends not only on fundamental but also on the level of aggregate investment, which reflects complementarity among investment decisions. Therefore, both variations in fundamental as well as non-fundamental sentiment can play a role. Individual investor increases his investment when the aggregate investment is higher. Such externalities in payoff are formalized as in Angeletos and Pavan (2007). Our model setup, however, features a three-way coordination among agents in the financial market: (1) coordination among investors; (2) coordination among analysts; and (3) coordination between analysts and investors.

[2]For more references, see Hirshleifer (2001) and Baker and Wurgler (2007) for survey treatment on this in the empirical finance literature.

Notably, sentiment arises in a self-fulfilling manner through information production and tranmission. Suppose that all analysts believe that aggregate investment reflects sentimental fluctuations. In response, financial reports written by analysts contain information about sentiment. As a result, information available to investors is contaminated by sentiment, therefore, individual investment decisions and aggregate investment are correlated with sentiment. The analysts' beliefs are self-fulfilling, that is, aggregate investment indeed depends on sentiment as a result.

Main Results. In this paper, we focus on symmetric-linear-Gaussian equilibrium where all actions i.e., report drafting or investing, are linear in signals and the information structure remains Gaussian. We demonstrate that, in addition to fundamental equilibrium where aggregate investment only responds to fundamentals, there exist sentiment equilibrium, in which, other than exogenous fundamentals, endogenous non-fundamental aggregate uncertainty affects aggregate investment.

We derive the necessary and sufficient condition for the existence of sentiment equilibrium, which highlights the importance of three elements in driving sentiment fluctuations, i.e., (1) the strategic interactions between analysts and investors; (2) imperfect coordination of analysts and investors; and (3) strong enough correlation between investment decision and sentiment. The necessary and sufficient condition states that as long as analysts care so much about reporting the level of aggregate investment K or have sufficiently strong incentive to align own report with the other analysts' reports, sentiment equilibrium arises.

Too much of concerns about the aggregate investment or sufficiently strong incentives to align own report with the others' divert analysts focus away from fundamental. Instead, analysts respond to actions of other agents, namely aggregate investment and average opinion of all other analysts. *Strong desire to coordinate but inability to do it perfectly* creates the room for sentiment to kick in inducing correlated beliefs and actions in a self-fulfilling manner. Acting as a coordination device, self-fulfilling sentiment creates co-movements in agents' beliefs and hence actions, which translates into non-fundamental volatility.

We demonstrate that the necessary and sufficient condition for the existence of sentiment equilibrium, ensures uniqueness in the distribution of sentiment too. Therefore, allocations in any sentiment equilibrium are isomorphic to the standard beauty contest with a common noise in the signal about fundamental.

In a sense, our model provides a micro-foundation for the common noises in the public information, which has been assumed exogenous in much of the literature.

Finally, considering the welfare implications of sentiment fluctuations, we demonstrate that sentiment fluctuations make the financial market more volatile by inducing non-fundamental volatility and by deteriorating the informativeness of investors' private information about fundamentals. However, sentiment, acting as a coordination device that helps agents to coordinate their beliefs and actions also reduces cross-sectional dispersion of investment, which is welfare improving. Therefore, whether sentiment deteriorates or enhances social welfare depends on the relative importance of volatility and cross-sectional dispersion in the determination of welfare loss. In other words, it is not necessarily true that sentiment is welfare reducing.

Relations to Literature. Our paper can be related to several strands of literature. First of all, our paper builds on Benhabib, Wang, and Wen (2015) and Benhabib, Liu, and Wang (2016) in generating self-fulfilling sentiment fluctuations. Benhabib, Wang, and Wen (2015) demonstrate that within a business cycle model that features Dixit-Stiglitz monopolistic competition, self-fulfilling sentiment may arise if firms cannot tell apart the idiosyncratic demand from the aggregate demand. In our paper, instead of focusing on the business cycle implications of sentiment fluctuations, we focus on the information production and transmission process within the financial markets and demonstrate that strategic interactions combined with imperfect coordination, create self-fulfilling sentiment fluctuations. Furthermore, in Benhabib, Liu, and Wang (2016), the authors focus on the "two-way feedback" between the real and financial markets and demonstrate that endogenous price signals in the financial market induce sentiment fluctuations in the real economy. The feedback creates multiple equilibria under complete information, which at the core of driving the continuum of sentiment equilibria differing in the variance of sentiment. While in our paper, the focus is the information production and transmission and the economy features unique equilibrium under complete information. What drives sentiment fluctuation is the strategic interactions and imperfect coordination instead of multiplicity in complete information allocation. Our paper also sheds light on the welfare implications of sentiment fluctuations, which is absent from either Benhabib, Wang, and Wen (2015) or Benhabib, Liu, and Wang (2016).

Secondly, our paper also builds on the literature of sender-receiver game in the modelling of information transmission process. Most of works in sender-receiver game literature assume that senders have perfect information over the fundamentals and focus on the strategic information disclosure among senders. See Takahashi and Ambrus (2008), Battaglini (2002) Bhattacharya and Mukherjee (2013), and Chan and Suen (2009) for example. Our paper contributes to this line of literature by demonstrating that when senders have incomplete information over fundamentals, there can be self-fulfilling sentiment fluctuations. To note that incomplete information is not only a more reasonable setup but also the necessary condition for the emergence of self-fulfilling sentiments. Chen and Suen (2018) also studies a sender-receiver game with senders having incomplete information over fundamental. We differ from them by introducing strategic complementarity among senders, which turns out to be the origins of sentiment fluctuations.

Thirdly, our paper directly relates to the literature of coordination games with incomplete information. See Morris and Shin (2002), Angeletos and Pavan (2007) and Myatt and Wallace (2012) in the context of abstract game theoretic models and Lorenzoni (2009), Angeletos and La'O (2013) and Woodford (2003) in the context of business cycle model. All of these works feature some exogenous public information over aggregate fundamental. Our model contributes to this line of literature by providing a micro-foundation for the public noise. If information were to be endogenously determined through the sender-receiver structure, the arising of sentiment fluctuations act as a public noise in the information available to receivers.

Finally, in the broader context, our paper also relates to the sunspot literature. See Diamond and Dybvig (1983), Benhabib and Farmer (1994) for example and Benhabib and Farmer (1999) for a survey treatment on this line of literature. In these works, non-convexity or strong form of complementarity generates multiple equilibria, each of whom differing in aggregate variables' response to fundamental. Multiple equilibria also arise in our paper. We differ from sunspot literature both in the nature and cause of the multiplicity. In our paper, multiple equilibria arise due to strategic interaction and imperfect coordination and multiplicity is over the realization of self-fulfilling sentiment. While, in case of the standard sunspot literature, multiplicity is about different responsiveness of endogenous variables to fundamentals and the cause of multiplicity is either non-convexity or strong strategic complementarity.

Layout. The rest of the paper is organized as follows. Section 3.2 sets up the model and defines the rational expectation equilibrium. Section 3.3 characterizes the equilibrium and highlights the origins of sentiment in financial markets. In Section 3.4, welfare implications of sentiment is discussed.

3.2. The Model

3.2.1. Agents, Actions and Payoffs

The economy consists of a continuum of investors indexed by $j \in J \equiv [0,1]$, as well as a continuum of analysts indexed by $i \in I \equiv [0,1]$. Investors need to choose an investment $k_j \in \mathcal{R}$ to maximize his payoff:

$$u_j = U\left(k_j, K, \sigma_k, \theta\right),$$

where $\theta \in \mathcal{R}$ stands for the fundamental; $K \equiv \int_j k_j dj$ for the aggregate investment; and $\sigma_k \equiv \left[\int_j \left(k_j - K\right)^2 dj\right]^{1/2}$ denotes the cross-sectional dispersion of investors' investment. Following Angeletos and Pavan (2007), U is assumed to be of the following form:

$$U\left(k, K, \sigma_k, \theta\right) = \left(k, K, \theta\right) \boldsymbol{S}_u \left(k, K, \theta\right)' + \frac{1}{2} U_{\sigma_k \sigma_k} \sigma_k^2,$$

where $\boldsymbol{S}_u$ denotes a 3×3 matrix consisting of model primitives.

Assumption 3.1. $U_{k\theta} > 0$ and $U_{kK} > 0$

Payoff of investor j depends on the fundamental θ, which can be interpreted as, for example, the underlying productivity of the economy or the profitability of the firm that attracts investment. We assume that $U_{k\theta} > 0$, which captures the idea that the marginal return to investment increases in θ. Investment decisions are of strategic complementarity, that is, $U_{kK} > 0$. The cross-sectional dispersion of investments σ_k affects the individual investor's payoff but has no strategic impact.[3] Analyst i in each analyst drafts a report $y_i \in \mathcal{R}$ to maximize his payoff:

$$v_i = V\left(y_i, \overline{y}, \sigma_y, K, \sigma_k, \theta\right),$$

where $\overline{y} \equiv \int_i y_i di$ is the average forecast reported by all analysts and $\sigma_y \equiv \left[\int_i \left(y_i - \overline{y}\right)^2\right]^{1/2}$ denotes the cross-sectional dispersion of these forecats. The utility function of

[3]The assumption that σ_z^2 has no strategic impact ensures certainty equivalence, which is the feature of many macro models using log-linearization.

analysts V is assumed to be of the following quadratic form:

$$V\left(y_i,\overline{y},\sigma_y,K,\sigma_k,\theta\right)=\left(y_i,\overline{y},K,\theta\right)\boldsymbol{S}_v\left(y_i,\overline{y},K,\theta\right)'+\left(\sigma_y,\sigma_k\right)\boldsymbol{S}_\sigma\left(\sigma_y,\sigma_k\right)',$$

where $\boldsymbol{S}_v$ is a 4×4 matrix and $\boldsymbol{S}_\sigma$ is a 2×2 matrix which consist of model primitives.

Assumption 3.2. $V_{y\theta}>0$, $V_{yK}>0$, $V_{y\overline{y}}>0$, and $V_{yy}<0$.

The fact that the analysts tend to herd their information for career concern has been documented empirically in numerous studies for example Hong, Kubik, and Solomon (2000). In our model, such an incentive is captured by $V_{y\overline{y}}>0$, which implies that individual analyst has incentive to benchmark his own forecast to the prevailing forecast of others'. We further assume that $V_{y\theta}>0$ and $V_{yK}>0$, so that analyst i's forecast is more sanguine, if either the fundamental θ or the aggregate investment K is higher. For example, analysts intend to offer forecasts that are close to the fundamental θ (e.g., the value of the firm whose security is being traded) as well as aggregate action K (e.g., other investors' investment decisions). Neither cross-sectional dispersions of information σ_y, nor that of investment decisions σ_k has any strategic role.

Assumption 3.3. $U_{kk}<0$, $-U_{kK}/U_{kk}<1$, $V_{yy}<0$ and $-V_{y\overline{y}}/V_{yy}<1$.

We further impose Assumption 3.3 so that equilibrium uniqueness is ensured in a useful benchmark case where there exists the strong complementarity among investors but the information structure is exogenous. Observe first that $V_{yy}<0$ and $U_{kk}<0$ ensure concavity in payoff of analysts and investors. Further, $-U_{kK}/U_{kk}<1$ and $-V_{y\overline{y}}/V_{yy}<1$ ensure a weak form of strategic complementarity following Angeletos and Lian (2016). The two sets of assumptions combined lead to the equilibrium uniqueness in the benchmark case.[4]

One illustrating example of the payoff functions is given as follows. As in Morris and Shin (2002), payoff of investors can take the following form:

$$u_j=U\left(k_j,K,\theta\right)=-\left(1-\alpha\right)\left(k_j-\theta\right)^2-\alpha\left(k_j-K\right)^2,$$

where α measure the degree of coordination motive of mass investors. Payoff

[4]In our model, information is endogenous and multiplicity arises because of strategic interactions between senders and receivers rather than the strong complementarity among investors.

of analyst i can take the following form:

$$v_i = V\left(y_i, \overline{y}, K, \theta\right) = -\left(1-\psi-\phi\right)\left(y_i-\theta\right)^2 - \psi\left(y_i-K\right)^2 - \phi\left(y_i-\overline{y}\right)^2,$$

where analyst i writes his report to minimize a weighted average of the distance of his own forecast (1) from fundamental θ, (2) from aggregate investment K and (3) from the average forecasts of other analysts $\overline{y}$. In this example, analysts forecast the market condition consists of fundamental, aggregate investment and average forecast with the relative importance being parameterized by $(1-\psi-\phi)$, ψ and ϕ respectively.

3.2.2. Information Structure

There exist two types of aggregate uncertainty that contribute to the fluctuation of aggregate investment K in this model, that is, fundamental θ and non-fundamental uncertainty, z (interpreted as sentiment). The fundamental θ is drawn by Nature from a Gaussian distribution, that is, $\theta \sim N(0, \sigma_\theta^2)$, where σ_θ^2 is the variance. Sentiment is an aggregate shock that may endogenously arise and its distribution is determined as an equilibrium outcome. The payoffs of analysts' and investors' do not depend on the sentiment shock. But as we will elaborate later, sentiment impacts on the aggregate investment K in a self-fulfilling manner.

Analyst i cannot can observe the realization of fundamental θ, but they receive private information x_i about it:

$$x_i = \theta + \iota_i$$

where ι_i is normally distributed with mean 0 and variance σ_ι^2, independent of fundamental θ and i.i.d across all analysts. To maintain tractability, we assume that analysts can observe the the sentiment z perfectly, if it arises.

Investors learn the fundamental θ and sentiment z, through reading financial forecasts by analysts.[5] However, investors are constrained by their capacity of learning, so that they cannot read all the reports from all analysts. We assume that investor j randomly selects and reads a subset $\mathcal{B}_j \subseteq I$ of financial forecasts. The randomness in $\mathcal{B}_j$ implies that investors act *as if* they receive only one piece

[5]The assumption that investors can only learn from financial forecasts offered by analysts is only a simplification but not crucial for our model mechanisms. Allowing investors to have access to private forecasts does not change our results

of information s_j, such that

$$s_j = \overline{y} + \varepsilon_j$$

where ε_j is i.i.d normal with mean 0 and variance $\left(\frac{1}{\kappa} - 1\right)\sigma_\iota^2$ and uncorrelated with $\overline{y}$. The parameter σ_ι^2 represents the heterogeneity across forecasts of analysts. The parameter $\kappa \in [0,1]$ characterizes the learning capacity of investors. If $\kappa = 0$, investors have no capacity of learning (or $\mathcal{B}_j$ is an empty set) and s_j is a pure noise. If $\kappa = 1$, investor j has infinite amount of capacity (or $\mathcal{B}_j = I$) and observes $\overline{y}$.

Such an assumption on information acquisition deserves elaboration. First, we assume that each investor has access to a subset of forecasts, given his capacity of learning, so that the informational heterogeneity across investors is maintained. Similar modelling device has been employed by Lorenzoni (2009).[6] Second, if sentiment does not arise, $\overline{y} = \theta$. Our assumption prevents the perfect information revealing, that is, investors acquire fundamental θ.

3.2.3. Equilibrium Definition

Definition 3.1 (Rational Expectation Equilibrium)**.** *A rational expectation equilibrium consists of (1) a set of analysts' information providing strategies $\{y_i^*(x_i,z)\}_{i\in I}$; (2) a set of investors' investment strategies $\left\{k_j^*(s_j)\right\}_{j\in J}$; (3) two aggregation functions for aggregate investment $K(\theta,z)$ and average forecasts $\overline{y}(\theta,z)$; and (4) an endogenous distribution P of sentiment z, which is independent of fundamental θ such that*

- *given $K(\theta,z)$ and $\overline{y}(\theta,z)$, $\mathcal{I}_i^{IB}$ and all other analysts' reports $\{y_m(x_m,z)\}_{m\neq i}$, analyst i maximizes own expected utility, i.e.,*

$$y_i^*(x_i,z) = \arg\max_{y_i\in\mathcal{R}} E\left[V\left(y_i,\overline{y},\sigma_y,K,\sigma_k,\theta\right)|\mathcal{I}_i^{IB}\right] \quad \forall i \in I \tag{3.1}$$

- *given $\{y_i(x_i,z)\}_{i\in I}$, $\mathcal{I}_i^{IB}$ and all the other investors' strategies $\{k_m(s_m)\}_{m\neq i}$,*

[6]Lorenzoni (2009) studies an island with dispersed information. In equilibrium, price of island j goods is a linear combination of aggregate TFP shock and island-specific TFP. If consumer on island i can trade with all other islands, he knows the aggregate TFP perfectly, which completely removes private information from the economy. Instead, if consumers are randomly assigned to a subset of goods for consumption hence only observe the prices of one subset of goods, it is as if each consumer receives another private signal about aggregate TFP. Therefore, information remains to be dispersed, even when trading across islands is allowed.

investor j maximizes own expected utility, i.e.,

$$k_j^*\left(s_j\right) = \underset{k_j \in \mathcal{R}}{\arg\max} E\left[U\left(k_j, K, \sigma_k, \theta\right) | \mathcal{I}_j^{IB}\right] \quad \forall j \in J \tag{3.2}$$

- *and the endogenous distribution of sentiment $P(z)$ ensures the consistency, i.e.,*

$$K(\theta, z) = \int_j k_j^*\left(\overline{y} + \iota_j\right) dj \quad \forall (\theta, z) \in \mathcal{R} \times \mathcal{R} \tag{3.3}$$

and

$$\overline{y}(\theta, z) = \int_i y_i^*\left(x_i, z\right) di \quad \forall (\theta, z) \in \mathcal{R} \times \mathcal{R} \tag{3.4}$$

In equilibrium, all investors and analysts believe that sentiment is drawn from a distribution P, aggregate investment is a mapping from both fundamental and sentiment $K(\theta, z)$ and average forecast is a mapping $\overline{y}(\theta, z)$. Analyst i maximizes his expected payoff, taking as given the reporting strategies of all other analysts, investing strategies of investors, the aggregate investment as well as the average forecast. Investor j maximizes his expected payoff, taking as given the investing strategies of all other investors as well as the reporting strategies of all analysts. The resulting decisions of analysts and investors are consistent with aggregate investment $K(\theta, z)$ and average forecast $\overline{y}(\theta, z)$. The distribution of sentiment P is endogenously determined to ensure the consistency.[7]

Observe that it is always possible that sentiment z has a degenerate distribution, in which $z = 0$ with probability 1. This corresponds to the case where all endogenous variables are solely driven by exogenous fundamental uncertainty θ. We label this case fundamental equilibrium. Otherwise, it is a sentiment equilibrium, when there exists non-degenerate sentiment.

3.3. Equilibrium Characterization

3.3.1. The Complete Information Benchmark

The model is defined to feature complete information if $\sigma_\iota = 0$, where analysts observe the fundamental θ or their private information is infinitely precise. In

[7] We exclude cross-sectional dispersions of financial reports σ_y and of investments σ_k from the equilibrium definition because they are assumed to be pure payoff externalities without any strategic roles in determining either optimal reporting strategy $y^*(x_i, z)$ or optimal investment strategy $k^*\left(s_j\right)$.

this case, analysts have common knowledge about θ and therefore provide homogenous forecasts. As a result, investors receive identical reports and the variance of ε_j varnishes to 0.

Proposition 3.1 (Complete Information). *Under complete information, there exists a unique equilibrium such that analyst i's forecast and investor j's investment are both linear in θ, i.e., $y_i = \overline{y} = \beta(\theta) \equiv \beta_0 + \beta_\theta \theta$ for all $i \in I$ and $k_j = K = \gamma(\theta) \equiv \gamma_0 + \gamma_\theta \theta$ for all $j \in J$ where*

$$\gamma_0 = \frac{-U_k(0,0,0,0)}{U_{kk} + U_{kK}} \qquad \gamma_\theta = \frac{-U_{k\theta}}{U_{kk} + U_{kK}} > 0$$

and

$$\beta_0 = \frac{-V_y(0,0,0,0,0,0) - V_{yK}\gamma_0}{V_{yy} + V_{y\overline{y}}} \qquad \beta_\theta = \frac{-V_{y\theta} - V_{yK}\gamma_\theta}{V_{yy} + V_{y\overline{y}}} > 0$$

Further, sentiment z has a degenerate distribution such that $z = 0$ with probability 1.

In the complete information benchmark, both forecasts of analysts and investments of investors increases with fundamental θ, i.e., $\beta_\theta, \gamma_\theta > 0$. Whenever fundamental θ is stronger, analysts write more positive report and investors invest more aggressively. Note that, when information is complete, the economy features no sentiment fluctuations, because the shared common knowledge among analysts and agents in the economy completely shuts down coordination frictions. Therefore, aggregate investment and average forecasts of analysts are also common knowledge. Recall that forecasts of analysts contain information about (1) fundamental θ, (2) aggregate investment K and (3) average forecasts of all analysts. As a result, investors can perfectly infer the fundamental θ. The whole economy features not only complete information, but also perfect information about fundamental θ.[8] Therefore, the complete information benchmark features perfect information transmission from analysts to investors. In equilibrium, beliefs are actions are anchored by fundamental θ, which leaves no room for any non-fundamental volatility due to sentiment fluctuations.

In what follows, to economize notations, without loss of generality, we normalize the complete information allocation such that $\gamma_0 = \beta_0 = 0$ and $\gamma_\theta = 1$

[8] We differentiate between notions of complete information and perfect information according to Angeletos and Lian (2016). Complete information means the existence of no private information or simply assuming common knowledge. While perfect information refers to the situation of knowing states of the economy perfectly. It is not necessarily true that one implies the other, depending on the assumptions of states and information structure.

by the imposing restrictions over model primitives. Upon normalization, under complete information, analyst i's forecast $y_i = \beta_\theta \theta$ and investor j's decision $k_j = \theta$.

Assumption 3.4. (Normalization). Payoff functions U and V are normalized such that $U_k(0,0,0,0) = V_y(0,0,0,0,0,0) = 0$ and $U_{k\theta} = U_{kk} + U_{kK}$.

3.3.2. Fundamental Equilibrium and Sentiment Equilibrium

In both fundamental and sentiment equilibrium, analysts and investors maximize their expected payoff given the relevant information as well as the equilibrium distributions of fundamental θ and sentiment z. Below, we characterize the analysts' optimal reporting and investor's optimal investing strategies, taking as given the distribution of sentiment $P(z)$.

Proposition 3.2 (Optimality). *Let α, χ, ϕ and ψ be such that*

$$\alpha = -\frac{U_{kK}}{U_{kk}} > 0 \qquad \chi = -\frac{V_{y\theta}}{V_{yy}} > 0 \qquad \phi = -\frac{V_{y\bar{y}}}{V_{yy}} \in (0,1) \qquad \psi = -\frac{V_{yK}}{V_{yy}} > 0$$

For any distribution of sentiment P, degenerate or non-degenerate, analyst i's optimal reporting strategy is such that

$$y_i = E\left[\chi\theta + \phi\bar{y} + \psi K | x_i, z\right] \quad \forall i \tag{3.5}$$

and investor j's optimal investing strategy is such that

$$k_j = E\left[(1-\alpha)\theta + \alpha K | s_j\right] \quad \forall j \tag{3.6}$$

In any equilibrium, optimal investing strategy of investor j is a weighted average of his belief over fundamental θ and his belief over aggregate investment K. Conditional on s_j and the distribution of sentiment P, equation (3.6) constitutes a fixed point condition over $\{k_j : j \in J\}$. Investment decisions are complementary to each other: when all investors other than j invest more, it incentivizes investor j to invest more.[9]

In any equilibrium, financial report y_i issued by analyst i increases in fundamental θ, aggregate investment K and the average forecast of all analysts.

[9]Such a fixed point is a common feature for models feature coordination motives and continuum action space, e.g., Morris and Shin (2002), Angeletos and Pavan (2007) and Myatt and Wallace (2012) in the abstract game theoretic models and Lorenzoni (2009), Angeletos and La'O (2013) and Woodford (2003) in the context of business cycle models.

Analyst i's report truthfully reflects his belief over the general market condition characterized by the fundamental θ and aggregate investment K. The weights assigned by analyst i to beliefs over fundamental and aggregate investment are χ and ϕ, respectively. Observe that equation (3.5) constitutes another fixed point condition over $\{y_i : i \in I\}$. The fact that $\phi > 0$ ensures that forecasts of analysts are of strategic complementarity.

Two fixed point conditions (3.5) and (3.6) are not independent of each other. Fixed point condition (3.5) has to be solved by taking solution of fixed point condition (3.6), i.e. K, as given. Similarly, fixed point condition (3.6) has to be solved by taking solution of fixed point condition (3.5), i.e. $\overline{y}$, as given. By solving (3.5) and (3.6), we can solve the equilibrium for any exogenous distribution of sentiment $P(z)$.

Note that fundamental equilibrium is equivalent to the case with degenerate sentiment distribution. Therefore, solving fixed point conditions with degenerate P leads to the characterization of fundamental equilibrium.

Lemma 3.1. Fundamental Equilibrium There always exists symmetric linear equilibrium without non-fundamental sentiment shock, such that

$$y_i(x_i) = \pi_{x,F} x_i \qquad k_j(s_j) = \pi_{s,F} s_j \qquad K(\theta) = \pi_{\theta,F}\theta.$$

where the elasticities $(\pi_{x,F}, \pi_{s,F}, \pi_{\theta,F})$ are the solutions to the system of equations (), () and (). Finally, information available to investor j is also linear in θ, such that

$$s_j = \pi_{x,F}\theta + \varepsilon_j \tag{3.7}$$

Throughout this paper, we focus on symmetric-linear-Gaussian equilibrium.

With exogenous information structure, symmetry comes from the fact that all agents are the same ex-ante and linearity arises due to quadratic payoff and Gaussian shocks.[10] When information structure is endogenous, as in our setup, linearity and Gaussian distributions are self-fulfilling in equilibrium. Linearity of action rules imply all endogenous variables are normally distributed, including information produced by analysts. Gaussian endogenous information further leads to linearity in action rules.

[10]With general payoff function, linearity can be understood as the results of log-linearization around a known steady state allocation, an approach commonly used in macro and finance literature.

Following the same logic, we are interested in characterizing sentiment equilibrium where sentiment z is normally distributed with mean 0 and variance σ_z^2 and contributes linearly to aggregate investment[11]:

$$K(\theta, z) = \pi_{\theta,S}\theta + z. \tag{3.8}$$

The nature of the sentiment z is also self-fulfilling. Suppose that all analysts believe that sentiment z affects aggregate investment K as in (3.8). Then financial reports written by analysts contain information about both fundamental θ and sentiment z. Hence, information available to investor j, i.e., s_j must be correlated with sentiment z, which further implies that k_j is correlated with sentiment z for all $j \in J$. Therefore, upon aggregation, it confirms the initial belief that aggregate investment is a function of sentiment z. However, to sustain the self-fulfilling sentiment, internal consistency, which requires that that the initial belief over aggregate investment coincides with equilibrium outcome induced by individual optimal investment decisions, has to be respected. It eventually pins down the distribution of sentiment $P(z)$ or on σ_z^2.

Proposition 3.3 (Sentiment Equilibrium)**.** *Symmetric linear equilibrium with Gaussian sentiment can be characterized by a 4-tuple of elasticities $(\pi_{x,S}, \pi_{z,S}, \pi_{s,S}, \pi_{\theta,S})$ together with the variance of sentiment σ_z^2 such that allocations are given by*

$$y_i(x_i, z) = \pi_{x,S}x_i + \pi_{z,S}z \qquad k_j(s_j) = \pi_{s,S}s_j \qquad K(\theta, z) = \pi_{\theta,S}\theta + z$$

and sentiment z is normally distributed with mean 0 and variance σ_z^2. Furthermore, symmetric linear equilibrium with Gaussian sentiment exists if and only if

$$(1-\alpha)\left[\left(\frac{\psi}{1-\phi}\right) - \chi\left(\frac{\sigma_\theta^2}{\sigma_\iota^2}\right)\right]\chi\left(\frac{\sigma_\theta^2}{\sigma_\iota^2}\right)\sigma_\theta^2 - \left(\frac{1}{\kappa} - 1\right)\sigma_\iota^2 > 0 \tag{3.9}$$

The variance of sentiment $\sigma_z^2 > 0$ is uniquely determined in equilibrium.

Recall that sentiment z is defined to impact aggregate investment on a one-to-one basis. Internal consistency over sentiment z requires that

$$\pi_{s,S}\pi_{z,S} = \mathbf{1}\left\{\sigma_z^2 > 0\right\} \tag{3.10}$$

[11]Sentiment z is normalized to have mean 0. In general, it can have a non-zero mean. However, it can be reduced to our model by simply re-labelling sentiment by subtracting the mean. Also, sentiment z is assumed to contribute to aggregate investment one to one. In general, it can be the case that $K(\theta, z) = \pi_{\theta,S}\theta + \delta z'$. This is equivalent to our model by imposing the fact that in equilibrium $\sigma_z^2 = \delta^2\sigma_{z'}^2$.

How much sentiment correlated with aggregate investment must be equal to the multiplication of how much it affects analysts' report $\pi_{z,S}$ and how much investors use private information to make investment $\pi_{s,S}$. Similarly, internal consistency over fundamental θ implies that

$$\pi_{\theta,S} = \pi_{s,S}\pi_{x,S} \tag{3.11}$$

Furthermore, given the fact that realization of sentiment z is common knowledge among all analysts, optimal reporting strategy of analysts (3.5) implies that response of forecasts of analyst i equals to the weighted average of the impacts of sentiment on aggregate investment K and on other analysts' forecasts

$$\pi_{z,S} = \psi \cdot \mathbf{1}\left\{\sigma_z^2 > 0\right\} + \phi\pi_{z,S} \tag{3.12}$$

Moreover, analyst i's posterior belief over fundamental θ is $E\left[\theta|x_i\right] = \left(\frac{\sigma_\theta^2}{\sigma_\theta^2+\sigma_\iota^2}\right)x_i$. Therefore, use of private information of analyst i must satisfy

$$\pi_{x,S} = \left(\chi + \phi\pi_{x,S} + \psi\pi_{\theta,S}\right)\left(\frac{\sigma_\theta^2}{\sigma_\theta^2 + \sigma_\iota^2}\right) \tag{3.13}$$

where $\chi + \phi\pi_{x,S} + \psi\pi_{\theta,S}$ measures how much analysts care about reporting fundamental θ consisting of direct concerns χ and indirect concerns through aggregate investment K and average forecasts by all other analysts $\bar{y}$. These four conditions uniquely pin down the 4-tuple of elasticities independent of the variance of sentiment σ_z^2.

Finally, the variance of sentiment σ_z^2 is endogenously determined to ensure that it provides the "right" incentive to investors in the use of private information:

$$\pi_{s,S} = (1-\alpha)\underbrace{\frac{\pi_{x,S}\sigma_\theta^2}{\pi_{x,S}^2\sigma_\theta^2 + \pi_{z,S}^2\sigma_z^2 + \left(\frac{1}{\kappa}-1\right)\sigma_\iota^2}}_{\rho\left(s_j,\theta\right)} + \alpha\underbrace{\frac{\left(\pi_{x,S}^2\sigma_\theta^2 + \pi_{z,S}^2\sigma_z^2\right)/\pi_{z,S}}{\pi_{x,S}^2\sigma_\theta^2 + \pi_{z,S}^2\sigma_z^2 + \left(\frac{1}{\kappa}-1\right)\sigma_\iota^2}}_{\rho\left(s_j,K\right)} \tag{3.14}$$

There are two opposing forces associated with a more volatile sentiment. Higher σ_z^2 contaminates the informativeness of s_j on fundamental θ, i.e. $\partial\rho\left(s_j,\theta\right)/\partial\sigma_z^2 < 0$. While, higher σ_z^2 implies that s_j convey more information over aggregate investment K, i.e., $\partial\rho\left(s_j,K\right)/\partial\sigma_z^2 > 0$.

Observe that σ_z^2 cannot be too large. Once it is i.e., $\sigma_z^2 \to +\infty$, all investors know that their information s_j is contaminated by sentiment z so much that the information s_j becomes no value in forecasting fundamentals θ: $\rho\left(s_j, \theta\right) = 0$. s_j simply becomes a piece of perfect information over aggregate investment K: $\rho\left(s_j, K\right) = 1$. This is because when sentiment becomes infinite volatile, non-fundamental volatility completely dominates fundamental volatility in the variations of aggregate investment K. With s_j conveying no information about fundamental θ and perfect information over aggregate investment K, individual investment decision would responds to s_j or equivalently K by a factor of α. Weak form complementarity $\alpha \in (0,1)$ directly implies that both k_j and K must be 0 naively, contradicting with the initial belief that economy features huge amount sentiment fluctuations at the aggregate level. Reducing σ_z^2 would possibly restore internal consistency. However, there is a natural lower bound 0 for σ_z^2. Therefore, it is possible for some model primitives, there presents no sentiment fluctuations at all.

Proposition 3.3 demonstrates that the necessary and sufficient condition to ensure existence is given by (3.9). More importantly, the same condition also implies uniqueness in σ_z^2. Uniquely determined σ_z^2 implies that all sentiment equilibria only differ in the realization of sentiment z. Therefore, allocations in sentiment equilibrium are isomorphic to standard beauty contest model that features a common noise $z \sim N\left(0, \sigma_z^2\right)$ in the . In this sense, our model provides a micro-foundation for the common noises in the public information literature.

Our model setup abstracts away any non-convexity and strong complementarity in actions so as to focus on how sentiment arises in the process of information production and transmission. Three key elements lead to the emergence of sentiment in this model: (1) the strategic interactions between analysts and investors; (2) the imperfect coordination of analysts and investors; and (3) strong enough correlation between individual investment decision k_j and sentiment z.

Corollary 3.1. There exists no sentiment equilibrium if either analysts' coordination incentive ψ, or heterogeneity in analysts' information σ_ι or investor's learning capacity κ is sufficiently low.

First of all, with little incentive to report the aggregate investment K, i.e. $\psi = 0$, there exists no strategic interactions between analysts and investors or forecasts offered by analysts are independent of investing strategies of investors. All analysts simply play a standard beauty contest as in Morris and Shin (2002) such that $y_i = \chi E\left[\theta | \mathcal{I}_i^{IB}\right] + \phi E\left[\overline{y} | \mathcal{I}_i^{IB}\right]$. The unique equilibrium, therefore, is y_i being solely a function of fundamental θ. From the perspectives of investors, they behave as if they are playing another beauty contest $k_j = (1-\alpha)\, E\left[\theta|\mathcal{I}_j^I\right] + \alpha E\left[K|\mathcal{I}_j^I\right]$ with exogenous information over fundamental θ. Therefore, in equilibrium, all endogenous outcomes of the model including information, beliefs and actions are anchored to fundamental θ, leaving no room for sentiment fluctuations.

Secondly, if we were able to remove the heterogeneity from the information set of analysts', i.e., $\sigma_\iota \to 0$, the fundamental becomes common knowledge among analysts. This is exactly the complete information benchmark in Proposition 3.1 featuring perfect information revelation from analysts to investors. Perfect information over fundamental θ anchors all agents' beliefs and actions, leaving no room for sentiment fluctuations.

Finally, when learning capacity is sufficiently low $\kappa = 0$, investors would ignore s_j when making investment decisions. As a result, there lacks enough correlation between individual investment k_j and sentiment z, which violates the consistency that requires aggregate investment K respond to sentiment z positively.

Corollary 3.2. Strong enough strategic interactions between analysts and investors or strong desire to align report with other analysts' reports are sufficient for sentiment fluctuations:

(1) Fixing primitives $\left(\alpha,\phi,\chi,\sigma_\theta^2,\sigma_\iota^2,\sigma_\iota^2,\kappa\right) \in (0,1) \times \mathcal{R}_+^5 \times (0,1)$, there exists a threshold $\overline{\psi}$ as a function of $\left(\alpha,\phi,\chi,\sigma_\theta^2,\sigma_\iota^2,\sigma_\iota^2,\kappa\right)$ such that there exist sentiment equilibrium if and only if $\psi > \overline{\psi} > 0$;

(2) Fixing primitives $\left(\alpha,\psi,\chi,\sigma_\theta^2,\sigma_\iota^2,\sigma_\iota^2,\kappa\right) \in (0,1) \times \mathcal{R}_+^5 \times (0,1)$, there exists a threshold $\overline{\phi}$ as a function of $\left(\alpha,\psi,\chi,\sigma_\theta^2,\sigma_\iota^2,\sigma_\iota^2,\kappa\right)$ such that there exist sentiment equilibrium if and only if $\phi > \overline{\phi} > 0$.

Both thresholds increase in learning capacity.

$$\frac{\partial\overline{\psi}}{\partial\kappa} < 0; \qquad\qquad \frac{\partial\overline{\phi}}{\partial\kappa} < 0.$$

Strategic interactions. Corollary 3.2 states that when coordination between individual analyst and mass investors is imperfect, i.e., $\sigma_\iota > 0$ and $\alpha < 1$, strong enough strategic interactions between analysts and mass investors ensures sentiment fluctuations. This is because with strong enough strategic interaction

($\psi > \overline{\phi}$), analysts' reports contains a lot of information on sentiment z given the fact that z affects aggregate investment K on a fixed one-to-one basis in equilibrium. Therefore, information available to investor j would contain quite a lot information over sentiment. It increases the responsiveness of investor j's investment decision k_j to sentiment z, which eventually supports the existence of sentiment equilibrium.

Career concerns. Corollary 3.2 also states that high enough ϕ or strong enough career concerns of analysts ensures sentiment fluctuations in equilibrium. There are two opposing forces at work. On the one hand, larger incentives to align own report with the rest of others' provide less incentive in the use of private information x_i, which makes $\overline{y}$ less unpredictable. Given the fact that aggregate investment K is a function of $\overline{y}$ for sure, more career concern results in less imperfection in coordination between analysts and investors. On the other hand, more incentive to align reports also provides more incentive to include sentiment z into financial report y_i given the fact that analyst i understands the others' reports also contain sentiment z under sentiment equilibrium. As a result, reliance of average reports $\overline{y}$ on sentiment z is increased, which for sure raises reliance of individual investment k_j on sentiment z. It turns out that the latter force dominates the former. Therefore, strong enough career concerns create room for sentiment fluctuations.

Learning capacity. Finally, Corollary 3.2 highlights that higher κ raises the threshold for ψ and ϕ. When learning capacity κ increases, investors can read more financial reports and get more precise information over average report $\overline{y}$. Therefore, investors rely more on s_j to make investment decisions in stead of relying on the the common prior, which raises the responsiveness of investment k_j to sentiment z. It also increases responsiveness of aggregate investment K to $\overline{y}$ which makes aggregate investment K less predictable from the perspective of an individual analyst. Hence less perfect coordination between individual analysts and investors. Both forces help create sentiment fluctuations.

3.4. Welfare Implications of Sentimental Fluctuations

In this section, we compare social welfare under fundamental equilibrium and sentiment equilibrium.

Definition 3.2. *Social welfare in our model is defined to be the ex-ante expected payoff* *of investors*

$$SW \equiv \int_{(\theta,z)} \int_{s_j} U\left(k\left(s_j\right), K\left(\theta,z\right), \sigma_k\left(\theta,z\right), \theta\right) dP\left(s_j|\theta,z\right) dP\left(\theta,z\right).$$

Following Angeletos and Pavan (2007), we define social welfare to be ex-ante expected payoff of the mass investors and the hypothetical planner is assumed to have no interest in protecting payoff of analysts. Following the same derivations in Angeletos and Pavan (2007),[12] social welfare can be rewritten as:

$$SW = E\left[U\left(\gamma^*\left(\theta\right), \gamma^*\left(\theta\right)0, \theta\right)\right] - \mathcal{L}, \tag{3.15}$$

where $\gamma^*\left(\theta\right) \equiv -U_k\left(0,0,0,0\right) - U_K\left(0,0,0,0\right) - \frac{U_{k\theta}+U_{K\theta}}{U_{kk}+2U_{kK}+U_{KK}}\theta$ denotes the efficient investing strategies under complete information and $\mathcal{L}$ denotes the welfare loss due to volatility and dispersion

$$\mathcal{L} = -\frac{U_{kk}+2U_{kK}+U_{KK}}{2}\underbrace{E\left[\left(K-\gamma^*\left(\theta\right)\right)^2\right]}_{\text{Volatility}} - \frac{U_{kk}+U_{\sigma\sigma}}{2}\underbrace{E\left[\left(k-K\right)^2\right]}_{\text{Dispersion}} \tag{3.16}$$

Social welfare is the sum of welfare under efficient investing strategies $E\left[U\left(\gamma^*, \gamma^* 0, \theta\right)\right]$ and a welfare loss $\mathcal{L}$ consisting of volatility $E\left[\left(K-\gamma^*\left(\theta\right)\right)^2\right]$ and cross-sectional dispersion $E\left[\left(k-K\right)^2\right]$ of investments. Observe that sentiment does not affect the former. Therefore, we focus on the welfare loss $\mathcal{L}$ for the comparison.

We further impose the Assumption 3.5 on U, which ensures that higher volatility or dispersion reduces social welfare and the economy features no welfare loss under complete information.

Assumption 3.5. $U_{kk}+2U_{kK}+U_{KK} < 0$, $U_{kk}+U_{\sigma\sigma} < 0$ and $\gamma^*\left(\theta\right) = \theta$.

In general, there could be multiple fundamental equilibria in the a sender-receiver game, where equilibria differ with each other in the response of aggregate investment K to fundamental θ. To ensure a clear comparison between sentiment equilibrium and fundamental equilibrium, we only focus on the case where fundamental equilibrium can be uniquely determined.

Proposition 3.4 (Welfare Implications)**.** *Volatility is higher in sentiment equilibrium but cross-sectional dispersion of investment is lower than that in fundamental equilibrium.*

[12]See appendix of Angeletos and Pavan (2007) for detailed derivations.

To understand the economic intuition behind this result, it is useful to understand the role of self-fulfilling sentiment z in the financial market. First of all, it induces non-fundamental volatility, which raises volatility of the market. On the other hand, considering making forecasts over fundamental θ, sentiment z deteriorates the informativeness of investor j's private information s_j over fundamental θ. Both of the two roles of sentiment contribute to higher volatility in sentiment equilibrium. However, sentiment fluctuations help investors to coordinate their actions, given it is a common shock. As a result, cross-sectional dispersion of investment σ_k decreases.

It is commonly believed that sentiment fluctuations or non-fundamental volatilities are detrimental to social welfare, which is true if sentiment fluctuations only affect volatilities of the economy. As we have demonstrate in Section 3.3, information heterogeneity ($\sigma_\iota^2 \neq 0$) is the necessary condition for the emergence of sentiment. Therefore, sentiment fluctuations must be coupled with a corresponding change in cross-sectional dispersion of the economy. The opposite impacts of sentiment z on volatility and cross-sectional dispersion lead to an ambiguous implication on social welfare. In fact, social welfare can be either higher or lower in sentiment equilibrium depending on the relative importance of volatility and dispersion in social welfare loss. If cross-sectional dispersion is far more important than volatility, sentiment can be a welfare improving, given it coordinates agents on both sides of information producing and receiving.

3.5. Conclusion

We conclude that financial sentiment may originates from endogenous information production and transmission of the financial market in a self-fulfilling manner. Within a sender-receiver game that features endogenous information production and transmission, we derive the necessary and sufficient conditions for the arising of sentiment fluctuations. The condition demonstrates that whenever the senders of information such as security analysts cannot perfectly coordinate with receivers of information such as investors, strong enough concerns of actions of the others, either level of aggregate investment or average opinions of other analysts, result in endogenous non-fundamental aggregate uncertainty, namely sentiment.

Strong desire to coordinate but inability to do it perfect creates enough room for sentiment to induce correlated beliefs and actions. At the aggregate level, they into non-fundamental volatilities, which is welfare deteriorating. While at the cross-section, they result in lower dispersion in investment decisions, which is welfare improving. Welfare implications of self-fulfilling sentiment are ambiguous depending on the relative importance of volatility and cross-sectional dispersion in the determination of welfare loss.

Appendix of Chapter 3

Proof of Proposition 3.1. Under complete information, analysts share a common belief. So do investors. Therefore, it must be the case that $y_i = y$ for all $i \in I$ and $k_j = k$ for all $j \in J$. Hence we have that $\overline{y} = y$, $K = k$ and $\sigma_y = \sigma_k = 0$.

Consider the problem of analysts first. Since analysts have perfect information over θ and z (if any). Therefore, in the optimality, it must be the case that

$$V_y\left(y,y,0,k,0,\theta\right) = 0 \Leftrightarrow V_y\left(0,0,0,0,0,0\right) + \left(V_{yy} + V_{y\overline{y}}\right)y + V_{yK}k + V_{y\theta}\theta = 0 \tag{3.17}$$

Next, consider the problem of investors. All investors have the same information y. Common knowledge here implies that $K = k$ is also common knowledge. Therefore, in the optimality, it must be the case that

$$E\left[U_k\left(k,k,0,\theta\right)|y\right] = 0 \Leftrightarrow U_k\left(0,0,0,0\right) + \left(U_{kk} + U_{kK}\right)k + U_{k\theta}E\left[\theta|y\right] = 0 \tag{3.18}$$

Equation (3.18) directly implies that k is a function of y only. Plugging it back to equation (3.17) implies that y is a pure function of θ, i.e., there is perfect information revealing from analysts to investors. Then it must be the case that $E\left[\theta|y\right] = \theta$.

Finally, we arrive at a system of two linear equations

$$\begin{cases} V_y\left(0,0,0,0,0,0\right) + \left(V_{yy} + V_{y\overline{y}}\right)y + V_{yK}k + V_{y\theta}\theta = 0 \\ U_k\left(0,0,0,0\right) + \left(U_{kk} + U_{kK}\right)k + U_{k\theta}\theta = 0 \end{cases}$$

Solving it leads to the expressions in Proposition 3.1. ∎

Proof of Proposition 3.2. Concavity of U implies that at the optimal, it must be the case that

$$E\left[U_k\left(k_j,K,\sigma_k,\theta\right)|\mathcal{I}_j^I\right] = 0$$

It can be shown that

$$U_k\left(k_j,K,\sigma_k,\theta\right) = U_k\left(\gamma\left(\theta\right),\gamma\left(\theta\right),0,\theta\right) + U_{kk}\left(k_j - \gamma\left(\theta\right)\right) + U_{kK}\left(K - \gamma\left(\theta\right)\right)$$

Further using the fact that $U_k(\gamma(\theta),\gamma(\theta),0,\theta) = 0$, we have

$$E\left[U_{kk}\left(k_j - \gamma(\theta)\right) + U_{kK}\left(K - \gamma(\theta)\right)|\mathcal{I}_j^I\right] = 0$$

Therefore, we have that

$$k_j = (1-\alpha)\, E\left[\gamma_k(\theta)\,|\mathcal{I}_j^I\right] + \alpha E\left[K|\mathcal{I}_j^I\right] \tag{3.19}$$

Similarly, concavity in V implies that at the optimal, it must be the case that

$$E\left[V_y\left(y_i,\overline{y},\sigma_y,K,\sigma_k,\theta\right)|\mathcal{I}_i^{IB}\right] = 0$$

It can be shown that

$$\begin{aligned} V_y\left(y_i,\overline{y},\sigma_y,K,\sigma_k,\theta\right) =& V_y\left(\beta(\theta),\beta(\theta),0,\gamma(\theta),0,\theta\right) \\ &+ V_{yy}\left(y_i - \beta(\theta)\right) + V_{y\overline{y}}\left(\overline{y} - \beta(\theta)\right) + V_{yK}\left(K - \gamma(\theta)\right) \end{aligned}$$

Further using the fact that $V_y(\beta(\theta),\beta(\theta),0,\gamma(\theta),0,\theta) = 0$, we have

$$E\left[V_{yy}\left(y_i - \beta(\theta)\right) + V_{y\overline{y}}\left(\overline{y} - \beta(\theta)\right) + V_{yK}\left(K - \gamma(\theta)\right)|\mathcal{I}_i^{IB}\right] = 0$$

Therefore, we have that

$$y_i = (1-\phi-\gamma_\theta\psi)\, E\left[\gamma_y(\theta)\,|\mathcal{I}_i^{IB}\right] + \phi E\left[\overline{y}|\mathcal{I}_i^{IB}\right] + \psi E\left[K|\mathcal{I}_i^{IB}\right] \tag{3.20}$$

Imposing the normalization of complete information equilibrium, i.e., $\gamma(\theta) = \gamma_y(\theta) = \theta$, leads to the expression in Proposition 3.2. ∎

Proof of Lemma 3.1. If there presents no sentiment in equilibrium and the equilibrium allocation is linear and symmetric, it must be the case that both $\overline{y}$ and K are linear in fundamental θ:

$$\overline{y} = \pi_{\overline{y}\theta}\theta \qquad\qquad K = \pi_{\theta,F}\theta$$

Then the information available to invest j would be

$$s_j = \pi_{\overline{y}\theta}\theta + \varepsilon_j$$

Following (3.6), solving the signal extraction problem leads to

$$k_j = \frac{\left[(1-\alpha)+\alpha\pi_{\theta,F}\right]\pi_{\overline{y}\theta}\sigma_\theta^2}{\pi_{\overline{y}\theta}^2\sigma_\theta^2+\left(\frac{1}{\kappa}-1\right)\sigma_\iota^2}s_j \equiv \pi_{s,F}s_j \tag{3.21}$$

Furthermore, solving the signal extraction problem of (3.5) leads to

$$y_i = \left(\chi+\phi\pi_{\overline{y}\theta}+\psi\pi_{\theta,F}\right)\left(\frac{\sigma_\theta^2}{\sigma_\theta^2+\sigma_\iota^2}\right)x_i \equiv \pi_{x,F}x_i \tag{3.22}$$

Finally, using the fact that $\pi_{\overline{y}\theta} = \pi_{x,F}$ and $\pi_{\theta,F} = \pi_{s,F}\pi_{\overline{y}\theta}$, we can arrive at the equilibrium conditions regulating $(\pi_{x,F}, \pi_{s,F}, \pi_{\theta,F})$:

$$\pi_{x,F} = (\chi+\phi\pi_{x,F}+\psi\pi_{\theta,F})\left(\frac{\sigma_\theta^2}{\sigma_\theta^2+\sigma_\iota^2}\right) \tag{3.23}$$

$$\pi_{s,F} = \frac{\left[(1-\alpha)+\alpha\pi_{\theta,F}\right]\pi_{x,F}\sigma_\theta^2}{\pi_{x,F}^2\sigma_\theta^2+\left(\frac{1}{\kappa}-1\right)\sigma_\iota^2} \tag{3.24}$$

and

$$\pi_{\theta,F} = \pi_{s,F}\pi_{x,F} \tag{3.25}$$

To show the existence, transform (3.23), (3.24) and (3.25) into one fixed point condition for $\pi_{\theta,F}$:

$$f(\pi_{\theta,F}) \equiv (1-\alpha)(1-\pi_{\theta,F})\left[\frac{\chi+\psi\pi_{\theta,F}}{1-\phi\left(\frac{\sigma_\theta^2}{\sigma_\theta^2+\sigma_\iota^2}\right)}\right]^2\left(\frac{\sigma_\theta^2}{\sigma_\iota^2+\sigma_\theta^2}\right)^2\sigma_\theta^2-\pi_{\theta,F}\left(\frac{1}{\kappa}-1\right)\sigma_\iota^2=0 \tag{3.26}$$

It is straight-forward to demonstrate existence since we have (1) $f(\pi_{\theta,F}) > 0$ for all $\pi_{\theta,F} < 0$ and (2) $f(\pi_{\theta,F}) < 0$ for all $\pi_{\theta,F} \geq 1$. ■

Proof of Proposition 3.3. When there presents sentiment, it must be the case that both $\overline{y}$ and K are given by:

$$\overline{y} = \pi_{\overline{y}\theta}\theta+\pi_{\overline{y}z}z \qquad K = \pi_{\theta,S}\theta+z$$

Then the information available to invest j would be

$$s_j = \pi_{\overline{y}\theta}\theta + \pi_{\overline{y}z}z + \varepsilon_j$$

Following (3.6), solving the signal extraction problem leads to

$$k_j = \frac{\left[(1-\alpha) + \alpha\pi_{\theta,S}\right]\pi_{\overline{y}\theta}\sigma_\theta^2 + \alpha\pi_{\overline{y}z}\sigma_z^2}{\pi_{\overline{y}\theta}^2\sigma_\theta^2 + \pi_{\overline{y}z}^2\sigma_z^2 + \left(\frac{1}{\kappa}-1\right)\sigma_\iota^2} s_j \equiv \pi_{s,S}s_j \tag{3.27}$$

Denote that $y_i = \pi_{x,S}x_i + \pi_{z,S}z$. Solving the signal extraction problem in (3.5) leads to

$$y_i = \left(\chi + \phi\pi_{\overline{y}\theta} + \psi\pi_{\theta,S}\right)\left(\frac{\sigma_\theta^2}{\sigma_\theta^2 + \sigma_\iota^2}\right)x_i + \left(\psi + \phi\pi_{z,S}\right)z \equiv \pi_{x,S}x_i + \pi_{z,S}z \tag{3.28}$$

The consistency condition over θ on K implies that

$$\pi_{\theta,S} = \pi_{s,S}\pi_{x,S} \tag{3.29}$$

Finally, consistency conditions over z on $\overline{y}$ and K imply that

$$\pi_{z,S} = \left(\frac{\psi}{1-\phi}\right)\mathbf{1}\left\{\sigma_z^2 > 0\right\} \tag{3.30}$$

$$\pi_{s,S}\pi_{z,S} = \mathbf{1}\left\{\sigma_z^2 > 0\right\} \tag{3.31}$$

Using the fact that $\pi_{\overline{y}\theta} = \pi_{x,S}$ and $\pi_{\theta,S} = \pi_{s,S}\pi_{\overline{y}\theta}$, we arrive at

$$\pi_{x,S} = \left(\chi + \phi\pi_{x,S} + \psi\pi_{\theta,S}\right)\left(\frac{\sigma_\theta^2}{\sigma_\theta^2 + \sigma_\iota^2}\right) \tag{3.32}$$

$$\pi_{\theta,S} = \pi_{s,S}\pi_{x,S} \tag{3.33}$$

$$\pi_{z,S} = \left(\frac{\psi}{1-\phi}\right)\mathbf{1}\left\{\sigma_z^2 > 0\right\} \tag{3.34}$$

$$\pi_{s,S}\pi_{z,S} = \mathbf{1}\left\{\sigma_z^2 > 0\right\} \tag{3.35}$$

and

$$\pi_{s,S} = \frac{\left[(1-\alpha) + \alpha\pi_{\theta,S}\right]\pi_{x,S}\sigma_\theta^2 + \alpha\pi_{z,S}\sigma_z^2}{\pi_{x,S}^2\sigma_\theta^2 + \pi_{z,S}^2\sigma_z^2 + \left(\frac{1}{\kappa}-1\right)\sigma_\iota^2} \tag{3.36}$$

Further using the fact that $\pi_{z,S}\pi_{\theta,S} = \pi_{x,S}$, we can obtain (3.14).

In the next step, solving (3.32), (3.33), (3.34) and (3.35) jointly leads to

$$\pi_{z,S} = \frac{\psi}{1-\phi} \qquad \pi_{s,S} = \frac{1-\phi}{\psi} \tag{3.37}$$

and

$$\pi_{x,S} = \chi\left(\frac{\sigma_\theta^2}{\sigma_\iota^2}\right) \qquad \pi_{\theta,S} = \left(\frac{1-\phi}{\psi}\right)\chi\left(\frac{\sigma_\theta^2}{\sigma_\iota^2}\right) \tag{3.38}$$

Plugging all these into (3.36), we arrive at a fixed point condition of σ_z^2:

$$g\left(\sigma_z^2\right) \equiv \frac{a_1\sigma_z^2 + b_1}{a_2\sigma_z^2 + b_2} - \frac{1-\phi}{\psi} = 0 \tag{3.39}$$

where we have that

$$a_1 = \alpha\pi_{z,S} \qquad b_1 = \left[(1-\alpha) + \alpha\pi_{\theta,S}\right]\pi_{x,S}\sigma_\theta^2 \tag{3.40}$$

$$a_2 = \pi_{z,S}^2 \qquad b_2 = \pi_{x,S}^2\sigma_\theta^2 + \left(\frac{1}{\kappa} - 1\right)\sigma_\iota^2 \tag{3.41}$$

$$g\left(\sigma_z^2\right) \equiv \frac{\left[(1-\alpha) + \alpha\pi_{\theta,S}\right]\pi_{x,S}\sigma_\theta^2 + \alpha\pi_{z,S}\sigma_z^2}{\pi_{x,S}^2\sigma_\theta^2 + \pi_{z,S}^2\sigma_z^2 + \left(\frac{1}{\kappa} - 1\right)\sigma_\iota^2} - \frac{1-\phi}{\psi} = 0 \tag{3.42}$$

It can be shown that

$$\frac{\partial g\left(\sigma_z^2\right)}{\partial\sigma_z^2} \propto \frac{a_1b_2 - a_2b_1}{\pi_{z,S}} = \alpha\left(\frac{1}{\kappa} - 1\right)\sigma_\iota^2 - (1-\alpha)\left(\frac{\psi}{1-\phi}\right)\pi_{x,S}\sigma_\theta^2 \tag{3.43}$$

Also, we have that

$$\lim_{\sigma_z^2 \to +\infty} g\left(\sigma_z^2\right) = -(1-\alpha)\left(\frac{1-\phi}{\psi}\right) \tag{3.44}$$

and

$$\lim_{\sigma_z^2 \to 0} g\left(\sigma_z^2\right) = \frac{b_1}{b_2} - \frac{1-\phi}{\psi} \tag{3.45}$$

$$= \frac{\left[(1-\alpha) + \alpha\pi_{\theta,S}\right]\pi_{x,S}\sigma_\theta^2}{\pi_{x,S}^2\sigma_\theta^2 + \left(\frac{1}{\kappa} - 1\right)\sigma_\iota^2} - \frac{1-\phi}{\psi} \tag{3.46}$$

$$\propto (1-\alpha)\left[\left(\frac{\psi}{1-\phi}\right) - \pi_{x,S}\right]\pi_{x,S}\sigma_\theta^2 - \left(\frac{1}{\kappa} - 1\right)\sigma_\iota^2 \tag{3.47}$$

We separate the problem into two cases.

1. If $(1-\alpha)\left[\left(\frac{\psi}{1-\phi}\right)-\pi_{x,S}\right]\pi_{x,S}\sigma_\theta^2-\left(\frac{1}{\kappa}-1\right)\sigma_\iota^2>0$, then we have $\lim_{\sigma_z^2\to 0} g\left(\sigma_z^2\right)>0$, $\lim_{\sigma_z^2\to+\infty} g\left(\sigma_z^2\right)<0$ and finally $\frac{\partial g\left(\sigma_z^2\right)}{\partial\sigma_z^2}<0$. This proves the uniqueness and existence of $\sigma_z^2\in(0,+\infty)$.

2. If $(1-\alpha)\left[\left(\frac{\psi}{1-\phi}\right)-\pi_{x,S}\right]\pi_{x,S}\sigma_\theta^2-\left(\frac{1}{\kappa}-1\right)\sigma_\iota^2<0$, then we have $\lim_{\sigma_z^2\to 0} g\left(\sigma_z^2\right)<0$, $\lim_{\sigma_z^2\to+\infty} g\left(\sigma_z^2\right)<0$ and finally $\frac{\partial g\left(\sigma_z^2\right)}{\partial\sigma_z^2}$ is either positive or negative with no critical point. Therefore, it must be the case that $g\left(\sigma_z^2\right)<0$ for all $\sigma_z^2>0$.

In sum, $(1-\alpha)\left[\left(\frac{\psi}{1-\phi}\right)-\pi_{x,S}\right]\pi_{x,S}\sigma_\theta^2-\left(\frac{1}{\kappa}-1\right)\sigma_\iota^2>0$ is the necessary and sufficient condition for $g\left(\sigma_z^2\right)=0$ to have a unique solution in the range of $(0,+\infty)$. ■

Proof of Corollary 3.1. Straight-forward following the (3.9). ■

Proof of Corollary 3.2. The existence of thresholds $\overline{\psi}$ and $\overline{\phi}$ directly follows Proposition 3.3. $\overline{\psi}$ can be proved to be decreasing in κ following the fact that (3.9) is increasing in both ψ and κ. The same argument applies to the comparative static analysis of $\overline{\phi}$. ■

Proof of Proposition 3.4. Under the condition (3.9), it can be shown that the fixed point condition (3.26) must be positive when evaluated at $\pi_{\theta,S}$, i.e., $f\left(\pi_{\theta,s}\right)>0$. Also note that if there presents a unique fundamental equilibrium, $f\left(\pi_{\theta,F}\right)$ must be decreasing when crossing the x-axis. Therefore, $f\left(\pi_{\theta,s}\right)>0$ directly implies that $\pi_{\theta,S}<\pi_{\theta,F}$. That is to say in fundamental equilibrium, aggregate investment K is more responsive to fundamental θ. Observe that in equilibrium both $\pi_{\theta,S}$ and $\pi_{\theta,F}$ are less than one. Then it must be the case that

$$Volatility_S\equiv E\left[\left(\pi_{\theta,S}\theta+z-\theta\right)^2\right]>E\left[\left(\pi_{\theta,F}\theta-\theta\right)^2\right]\equiv Volatility_F$$

Also, it can be shown that for $T\in\{F,S\}$ we have that

$$1=\left(\chi\frac{1}{\pi_{x,T}}+\phi+\psi\pi_{s,T}\right)\left(\frac{\sigma_\theta^2}{\sigma_\theta^2+\sigma_\iota^2}\right)$$

where $\pi_{x,T}=\frac{\chi+\psi\pi_{\theta,T}}{1-\phi\left(\frac{\sigma_\theta^2}{\sigma_\theta^2+\sigma_\iota^2}\right)}\left(\frac{\sigma_\theta^2}{\sigma_\theta^2+\sigma_\iota^2}\right)$. Therefore, it is straight-forward to show that

$$\pi_{s,S}<\pi_{s,F} \tag{3.48}$$

Finally, given the fact that cross-sectional dispersion in investments is $\pi_{s,S}^2\left(\frac{1}{\kappa}-1\right)\sigma_\iota^2$ under sentiment equilibrium and is $\pi_{s,F}^2\left(\frac{1}{\kappa}-1\right)\sigma_\iota^2$ under fundamental equilibrium. It is straight-forward to show that sentiment equilibrium features lower cross-sectional dispersion. ■

Bibliography

Angeletos, G.-M., F. Collard, and H. Dellas (2016a). An anatomy of the business cycle data. *Working Paper*.

Angeletos, G.-M., F. Collard, and H. Dellas (2016b). Quantifying confidence. *Working Paper*.

Angeletos, G.-M., L. Iovino, and J. La'O (2016). Real rigidity, nominal rigidity, and the social value of information. *American Economic Review 106*(1), 200–227.

Angeletos, G.-M. and J. La'O (2009). Noisy business cycles. *NBER Macroeconomics 2009*.

Angeletos, G.-M. and J. La'O (2013). Sentiments. *Econometrica 81*(2), 739–779.

Angeletos, G.-M. and C. Lian (2016). Chapter 14 incomplete information in macroeconomics: Accommodating frictions in coordination. *Handbook of Macroeconomics Volume 2 (Supplement C)*, 1065 – 1240.

Angeletos, G.-M. and A. Pavan (2006). Socially optimal coordination: Characterization and policy implication. *Journal of the European Economic Association 5(2-3)*, pp. 585–593.

Angeletos, G.-M. and A. Pavan (2007). Efficient use of information and social value of information. *Econometrica 75*(4), 1103–1142.

Baker, M. and J. Wurgler (2007). Investor sentiment in the stock market. *Journal of Economic Perspectives 21*(2), 129–152.

Baqaee, D. R. (2017). Asymmetric inflation expectations, downward rigidity of wages, and asymmetric business cycles. *Working Paper*.

Barsky, R. B. and E. R. Sims (2012). Information, animal spirits, and the meaning of innovations in consumer confidence. *American Economic Review 102*(4), 1343–77.

Basu, S., J. G. Fernald, and M. S. Kimball (2006). Are Technology Improvements Contractionary? *American Economic Review 96*(5), 1418–1448.

Battaglini, M. (2002). Multiple referrals and multidimensional cheap talk. *Econometrica 70*(4), 1379–1401.

Beaudry, P. and F. Portier (2004). An exploration into Pigou's theory of cycles. *Journal of Monetary Economics 51*(6), 1183–1216.

Beaudry, P. and F. Portier (2006). Stock prices, news, and economic fluctuations. *The American Economic Review 96*(4), 1293–1307.

Beaudry, P. and F. Portier (2014). Understanding noninflationary demand-driven business cycles. *NBER Macroeconomics Annual 28*(1), 69–130.

Benhabib, J. and R. E. Farmer (1994). Indeterminacy and increasing returns. *Journal of Economic Theory 63*(1), 19 – 41.

Benhabib, J. and R. E. Farmer (1999). Chapter 6 indeterminacy and sunspots in macroeconomics. Volume 1 of *Handbook of Macroeconomics*, pp. 387 – 448. Elsevier.

Benhabib, J., X. Liu, and P. Wang (2016). Sentiments, financial markets, and macroeconomic fluctuations. *Journal of Financial Economics 120*(2), 420 – 443.

Benhabib, J., P. Wang, and Y. Wen (2015). Sentiments and aggregate demand fluctuations. *Econometrica 83*(2), 549–585.

Bhandari, A., J. Borovicka, and P. Ho (2016). Identifying ambiguity shocks in business cycle models using survey data. *Working Paper*.

Bhattacharya, S. and A. Mukherjee (2013). Strategic information revelation when experts compete to influence. *The RAND Journal of Economics 44*(3), 522–544.

Bidder, R. and M. Smith (2012). Robust animal spirits. *Journal of Monetary Economics 59*(8), 738 – 750.

Blanchard, O. J., J.-P. L'Huillier, and G. Lorenzoni (2013). News, Noise, and Fluctuations: An Empirical Exploration. *American Economic Review 103*(7), 3045–3070.

Blanchard, O. J. and D. Quah (1989). The Dynamic Effects of Aggregate Demand and Supply Disturbances. *American Economic Review 79*(4), 655–673.

Bloom, N. (2009). The impact of uncertainty shocks. *Econometrica 77*(3), 623–685.

Bloom, N., M. Floetotto, N. Jaimovich, I. Saporta-Eksten, and S. Terry (2016). Really uncertain business cycles. *Working Paper*.

Boyarchenko, N. (2012). Ambiguity shifts and the 2007-2008 financial crisis. *Journal of Monetary Economics 59*(5), 493 – 507.

Chan, J. and W. Suen (2009). Media as watchdogs: The role of news media in electoral competition. *European Economic Review 53*(7), 799 – 814.

Chari, V. V., P. J. Kehoe, and E. R. McGrattan (2007). Business cycle accounting. *Econometrica 75*(3), 781–836.

Chen, H., Y. K. Lu, and W. Suen (2016). The power of whispers: A theory of rumor, communication, and revolution*. *International Economic Review 57*(1), 89–116.

Chen, H. and W. Suen (2016). Falling dominoes: A theory of rare events and crisis contagion. *American Economic Journal: Microeconomics 8*(1), 228–55.

Chen, H. and W. Suen (2018). Competition for attention in the news media market. *Working Paper*.

Colombo, L., G. Femminis, and A. Pavan (2012). Information acquisition and welfare. Technical report.

Cornand, C. and F. Heinemann (2008). Optimal degree of public information dissemination. *The Economic Journal 118*(528), pp. 718–742.

Diamond, D. W. and P. H. Dybvig (1983). Bank runs, deposit insurance, and liquidity. *Journal of Political Economy 91*(3), 401–419.

Epstein, L. G. and M. Schneider (2003). Recursive multiple priors. *113*, 1–31.

Epstein, L. G. and M. Schneider (2008). Ambiguity, information quality, and asset pricing. *The Journal of Finance 63*(1), 197–228.

Gali, J. (1999). Technology, Employment, and the Business Cycle: Do Technology Shocks Explain Aggregate Fluctuations? *American Economic Review 89*(1), 249–271.

Gilboa, I. and D. Schmeidler (1989). Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics 18*(2), 141–153.

Hall, R. (2001). Struggling to understand the stock market. *91*, 1–11.

Hanany, E. and P. Klibanoff (2009). Updating ambiguity averse preferences. *The B.E. Journal of Theoretical Economics 9*.

Hansen, L. P. and T. J. Sargent (2001a). Acknowledging misspecification in macroeconomic theory. *Review of Economic Dynamics 4*(3), 519 – 535.

Hansen, L. P. and T. J. Sargent (2001b). Robust control and model uncertainty. *American Economic Review 91*(2), 60–66.

Hellwig, C., S. Kohls, and L. Veldkamp (2012). Information choice technologies. *American Economic Review 102*(3), 35–40.

Hellwig, C. and L. Veldkamp (2009). Knowing what others know: Coordination motives in information acquisition. *Review of Economic Studies 76*(1), 223–251.

Hirshleifer, D. (2001). Investor psychology and asset pricing. *The Journal of Finance 56*(4), 1533–1597.

Hong, H., J. D. Kubik, and A. Solomon (2000). Security analysts' career concerns and herding of earnings forecasts. *The RAND Journal of Economics 31*(1), 121–144.

Huo, Z. and N. Takayama (2015). Higher order beliefs, confidence, and business cycles. *Working Paper*.

Ilut, C. (2012). Ambiguity aversion: Implications for the uncovered interest rate parity puzzle. *American Economic Journal: Macroeconomics 4*(3), 33–65.

Ilut, C., M. Kehrig, and M. Schneider (2016). Slow to hire, quick to fire: Employment dynamics with asymmetric responses to news. *Working Paper*.

Ilut, C. and H. Saijo (2016). Learning, confidence, and business cycles. *Working Paper*.

Ilut, C. and M. Schneider (2014). Ambiguous business cycles. *American Economic Review 104*(8), 2368–2399.

Jaimovich, N. and S. Rebelo (2009). Can News about the Future Drive the Business Cycle? *American Economic Review 99*(4), 1097–1118.

Kahneman, D. (1973). *Attention and Effort*. Englewood Cliffs, New Jersey: Prentice-Hall Inc.

Klibanoff, P., M. Marinacci, and S. Mukerji (2005). A smooth model of decision making under ambiguity. *Econometrica 73*(6), 1849–1892.

Klibanoff, P., M. Marinacci, and S. Mukerji (2009). Recursive smooth ambiguity preferences. *Journal of Economic Theory 144*(3), 930 – 976.

Lorenzoni, G. (2009). A theory of demand shocks. *American Economic Review 99*(5), 2050–84.

Luo, Y. (2008). Consumption dynamics under information processing constraints. *Review of Economic Dynamics 11*(2), 366–385.

Luo, Y. and E. R. Young (2010). Risk-sensitive consumption and savings under rational inattention. *American Economic Journal: Macroeconomics 2*(4), 281–325.

Maccheroni, F., M. Marinacci, and D. Ruffino (2013). Alpha as ambiguity: Robust mean-variance portfolio analysis. *Econometrica 81*(3), 1075–1113.

Maćkowiak, B. A. and M. Wiederholt (2009). Optimal sticky prices under rational inattention. *American Economic Review 99*(3).

Maćkowiak, B. A. and M. Wiederholt (2011). Inattention to rare events. CEPR Discussion Papers 8626, C.E.P.R. Discussion Papers.

Mavroeidis, S., M. Plagborg-Møller, and J. H. Stock (2014). Empirical Evidence on Inflation Expectations in the New Keynesian Phillips Curve. *Journal of Economic Literature 52*(1), 124–188.

Mondria, J. (2010). Portfolio choice, attention allocation, and price comovement. *Journal of Economic Theory 145*(5), 1837 – 1864.

Morris, S. and H. S. Shin (2002). Social value of public information. *American Economic Review 92*(5), 1521–1534.

Myatt, D. P. and C. Wallace (2009). On the sources and value of information: Public announcements and macroeconomic performance. Technical report.

Myatt, D. P. and C. Wallace (2012). Endogenous information acquisition in coordination games. *79*(1), 340–374.

Peng, L. (2005). Learning with information capacity constraints. *The Journal of Financial and Quantitative Analysis 40*(2), pp. 307–329.

Peng, L. and W. Xiong (2006). Investor attention, overconfidence and category learning. *Journal of Financial Economics 80*(3), 563 – 602.

Peter Hansen, L. and M. Marinacci (2016, 11). Ambiguity aversion and model misspecification: An economic perspective. *31*, 511–515.

Shiller, R. J. (2003). From efficient markets theory to behavioral finance. *Journal of Economic Perspectives 17*(1), 83–104.

Sims, C. A. (2003). Implications of rational inattention. *Journal of Monetary Economics 50*(3), 665–690.

Sims, C. A. (2005). Rational inattention: a research agenda. Technical report.

Sims, C. A. (2010). Rational inattention and monetary economics. In B. M. Friedman and M. Woodford (Eds.), *Handbook of Monetary Economics*, Volume 3 of *Handbook of Monetary Economics*, Chapter 4, pp. 155–181. Elsevier.

Svensson, L. E. O. (2006). Social value of public information: Comment: Morris and shin (2002) is actually pro-transparency, not con. *American Economic Review 96*(1).

Takahashi, S. and A. Ambrus (2008). The multi-sender cheap talk with restricted state spaces. *Theoretical Economics 3*.

Van Nieuwerburgh, S. and L. Veldkamp (2009). Information immobility and the home bias puzzle. *The Journal of Finance 64*(3), 1187–1215.

Van Nieuwerburgh, S. and L. Veldkamp (2010). Information acquisition and under-diversification. *The Review of Economic Studies 77*(2), 779–805.

Veldkamp, L. (2011). *Information choice in macroeconomics and finance / Laura L. Veld- kamp*. Princeton University Press, Princeton, N.J.

Verma, R. and P. Verma (2007). Noise trading and stock market volatility. *Journal of Multinational Financial Management 17*(3), 231 – 243.

Woodford, M. (2003). Imperfect common knowledge and the effects of monetary policy. *Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps*.

Yang, M. (2015). Coordination with flexible information acquisition. *Journal of Economic Theory 158, Part B*(0), 721 – 738.

Zhang, D. (2017). Term structure, forecast revision and the information channel of monetary policy. *Working Paper*.

GUANGYU PEI

Personal Information

Date of Birth	Place of Birth
22.02.1989	Jilin, China

Education

Sep. 2013 - Jul. 2018	Ph.D. Studies in Economics, University of Zurich, Switzerland
Feb. 2017 - Jun. 2017	Visiting Student, Department of Economics, MIT, USA
Sep. 2012 - Jul. 2013	Ph.D. Studies in Economics, Stockholm University, Sweden
Sep. 2009 - Jun. 2013	Bachelor in Economics and Finance, The University of Hong Kong, Hong Kong

Research Fields

Primary	Macroeconomics. Business Cycles. Information Economics.
Secondary	Coordination Games with Incomplete Information